Discovery and functional characterisation of novel neuropeptide signalling systems in echinoderms

Semmens, Dean Colin

Discovery and functional characterisation of novel neuropeptide signalling systems in echinoderms

Dean Colin Semmens

Submitted in partial fulfillment of the requirements of the Degree of Doctor of Philosophy

November 2015

Statement of originality

I, Dean Colin Semmens, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: Dean Colin Semmens

Date: November 2015

Details of collaboration and publications:

(1) Collaborations:

---

(1) Dr. Olivier Mirabeau

Cancer Genetics Unit, Institut Curie, Paris, France

Collaboration with Dr. Olivier Mirabeau for the identification of some of the transcripts comprising neuropeptide and peptide hormone precursors in the starfish *Asterias rubens* as reported in **chapter 2**.

(2) Prof. James H. Scrivens and Dr. Susan E. Slade

Waters/Warwick Centre for BioMedical Mass Spectrometry and Proteomics, School of Life Sciences, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK

Collaboration with Prof. James H. Scrivens and Susan E. Slade for mass spectrometric confirmation of a vasopressin/oxytocin-type peptide ("asterotocin") and an NG peptide ("NGFFYamide") in the starfish *A. rubens* as reported in **chapters 3 and 4** respectively.

(2) Publications:

Semmens, D. C., Dane, R. E., Pancholi, M. R., Slade, S. E., Scrivens, J. H. and Elphick, M. R. (2013). Discovery of a novel neurophysin-associated neuropeptide that triggers cardiac stomach contraction and retraction in starfish. *J Exp Biol* **216(21)**, 4047-4053. doi: 10.1242/jeb.092171.

Elphick, M. R., **Semmens, D. C**., Blowes, L. M., Levine, J., Lowe, C. J., Arnone, M. I. and Clark, M. S. (2015). Reconstructing SALMFamide neuropeptide precursor evolution in the phylum Echinodermata: ophiuroid and crinoid sequence data provide new insights. *Front Endocrinol* **6(2)**. doi: 10.3389/fendo.2015.00002.

Semmens, D. C., Beets, I., Rowe, M. L., Blowes, L. M., Oliveri, P. and Elphick, M. R. (2015). Discovery of sea urchin NGFFFamide receptor unifies a bilaterian neuropeptide family. *Open Biol* **5**, 150030. doi: 10.1098/rsob.150030.

Christensen, A. B., Herman, J. L., Elphick, M. R., Kober, K. M., Janies, D., Linchangco, G., **Semmens, D. C.**, Bailly, X., Vinogradov, S. N., Hoogewijs, D. (2015). Phylogeny of echinoderm hemoglobins. *PLoS ONE* **10(8)**, e0129668. doi: 10.1371/journal.pone.0129668.

Note: This paper contains *A. rubens* sequence data that is not reported in this thesis.

Jones, C. E., Zandawala, M., **Semmens, D. C.**, Anderson, S., Hanson, G. R., Janies, D. and Elphick, M. R. (2016). Identification of a neuropeptide precursor protein that gives rise to a "cocktail" of peptides that bind Cu(II) and generate metal-linked dimers. *Biochim Biophys Acta* **1860**, 57-66. doi: 10.1016/j.bbagen.2015.10.008.

Abstract

Neuropeptides are evolutionarily ancient mediators of neuronal signalling that regulate a diverse range of physiological processes and behaviours. Recent advances in comparative genomics/transcriptomics are providing opportunities to analyse neuropeptide systems in a wider range of phyla. The echinoderms (e.g. starfish, sea urchins and sea cucumbers) are of particular interest in this respect because as deuterostomian invertebrates they occupy an intermediate position in animal phylogeny, bridging model protostomian invertebrates (e.g. *Drosophila*) to the vertebrates. Herein, the common European starfish *Asterias rubens* has been established as a model system for neuropeptide research.

Generation and analysis of a neural transcriptome dataset from *A. rubens* has enabled the identification of thirty-five novel neuropeptide precursors and a number of candidate receptors. Interestingly, precursors of two kisspeptin-type peptides (the first identified in a non-chordate species), a melanin-concentrating hormone-type peptide (the first identified outside of the vertebrates) and two tachykinin-type peptides (the first identified in an ambulacrarian species) were identified. Discovery of this repertoire of neuropeptide precursors has provided the foundation for a comprehensive analysis of the physiological functions of neuropeptides in starfish.

Investigation of the physiological roles of two neuropeptides – a vasopressin/oxytocin-type peptide ("asterotocin") and an NG peptide ("NGFFYamide") – has revealed roles in the remarkable process of extraoral feeding in starfish. Discovery of neuropeptides that trigger cardiac stomach eversion (asterotocin) and retraction (NGFFYamide) provides a novel insight into the neural regulation of starfish feeding and a rationale for chemically based strategies to control starfish that feed on economically important shellfish or protected marine fauna.

Finally, characterisation of the receptor for the NG peptide NGFFFamide in the sea urchin *Strongylocentrotus purpuratus* has unified a bilaterian neuropeptide family that includes neuropeptide-S-type peptides in tetrapod vertebrates, NG peptides in deuterostomian invertebrates and crustacean cardioactive peptide-type peptides in protostomian invertebrates.

Table of contents

1. General introduction 30

1.1. Neuropeptides 30

1.1.1. Neuropeptide structural characteristics 31

1.1.2. Neuropeptide release 34

1.2. G-protein coupled receptors (GPCRs) 36

1.3. Bilaterian neuropeptide signalling systems 38

1.3.1. The Bilateria 38

1.3.2. The Bilateria: providing insights into the evolutionarily ancient origins of neuropeptide signalling systems 39

1.4. The echinoderms 41

1.4.1. The echinoderms: a model system for neuropeptide research 43

1.5. The common European starfish *A. rubens* 46

1.5.1. *A. rubens* external anatomy 46

1.5.2. *A. rubens* internal anatomy 47

1.5.3. *A. rubens*: a model system for neuropeptide research 51

1.6. Aims and objectives 53

2. Identification of transcripts encoding neuropeptide precursors and neuropeptide receptors in the starfish *A. rubens*

2.1. General introduction 55

2.1.1. The Ambulacraria 56

2.1.2. Ambulacrarian neuropeptide signalling systems: the hemichordates 56

2.1.3. Ambulacrarian neuropeptide signalling systems: the echinoderms 57

2.1.4. *A. rubens*: a model system for neuropeptide research 58

2.1.5. Aims and objectives 59

2.2. Methods 60

2.2.1. Animal collection 60

2.2.2. RNA isolation 60

2.2.3. Illumina Hi-Seq 60

2.2.4. Transcriptome assembly 61

2.2.5. Bioinformatic identification of novel neuropeptide/peptide hormone precursors 61

2.2.6. Multiple sequence alignments 63

2.2.7. Bioinformatic identification of candidate neuropeptide/peptide hormone receptors 65

2.2.8. Neighbour joining (NJ) trees 66

2.3. Results and Discussion 67

2.3.1. Precursors of neuropeptides that have been identified previously in *A. rubens* – the SALMFamides S1 and S2 67

2.3.2. Precursors of neuropeptides that belong to bilaterian neuropeptide families 75

2.3.2.1. VP/OT-type precursor (“asterotocin” precursor) 78

2.3.2.2. NGFFYamide precursor 79

2.3.2.3. GnRH-type precursor 1 (ArGnRHP1) 80

2.3.2.4. GnRH-type precursor 2 (ArGnRHP2) 84

2.3.2.5. TRH-like precursor (ArTRHLP) 86

2.3.2.6. TK-like precursor (ArTKLP) 90

2.3.2.7. CCK-like precursor (ArCCKLP) 94
2.3.2.8. OX-like precursors (ArOXLP1-2) 99
2.3.2.9. LQ-like precursor (ArLQLP) 104
2.3.2.10. KP-like precursor (ArKPLP) 108
2.3.2.11. SMS-like precursor (ArSMSLP) 112
2.3.2.12. MCH-like precursor (ArMCHLP) 116
2.3.2.13. CT-like precursor (ArCTLPP) 119
2.3.2.14. CRH-like precursor (ArCRHLP) 123
2.3.2.15. PDF-like precursor (ArPDFLP) 127
2.3.2.16. PP/orcokinin-like precursor 1 (ArPPLNP1) 130

2.3.3. Precursors of neuropeptides and peptide hormones that contain multiple cysteine (C) residues 134
2.3.3.1. GPA2/5-type precursors (ArGPA2-1-2 and ArGPB5-1-3) 135
2.3.3.2. Bursicon α/β-like precursors (ArBALP and ArBBLP) 139
2.3.3.3. GSS-like precursor (ArGSSLP) 142
2.3.3.4. Relaxin-like precursor (ArRLP) 144
2.3.3.5. IGF-like precursors (ArIGF1-2) 147

2.3.4. Precursors of putative neuropeptides that appear to be unrelated to known bilaterian neuropeptide families 152
2.3.4.1. AN peptide-type precursor (ArANPP) 153
2.3.4.2. Arnp11 155
2.3.4.3. Arnp15a-b 157
2.3.4.4. Arnp18 159

2.4. General discussion 160

3. Discovery and functional characterisation of a novel vasopressin/oxytocin (VP/OT)-type neuropeptide ("asterotocin") in the starfish *A. rubens*

3.1. Introduction 163

3.1.1. VP/OT-type neuropeptide function in the vertebrates 166

3.1.2. VP/OT-type neuropeptide function across the animal kingdom 166

3.1.2.1. VP/OT-type peptides in the deuterostomes 167

3.1.2.2. VP/OT-type peptides in the protostomes 168

3.1.3. Investigating VP/OT-type peptides in *A. rubens* 169

3.1.4. Aims and objectives 170

3.2. Methods 171

3.2.1. Bioinformatic identification of the asterotocin precursor and a candidate asterotocin receptor in *A. rubens* 171

3.2.2. Cloning and sequencing of the asterotocin precursor 172

3.2.3. Mass spectrometric confirmation of asterotocin 172

3.2.4. Localisation of the asterotocin precursor in *A. rubens* using mRNA *in situ* hybridisation (mRNA ISH) 174

3.2.4.1. Probe synthesis 174

3.2.4.2. Tissue fixation 175

3.2.4.3. Tissue sectioning 175

3.2.4.4. Probe hybridisation and immunodetection 176

3.2.5. Analysis of the *in vivo* effects of asterotocin in *A. rubens* 178

3.2.6. Analysis of the *in vitro* effects of asterotocin on cardiac stomach preparations from *A. rubens* 179

3.3. Results 181

3.3.1. Bioinformatic identification and cloning of the asterotocin precursor and identification of a candidate VP/OT-type receptor in *A. rubens* 181

3.3.2. Mass spectrometric confirmation of the structure of asterotocin 184

3.3.3. Expression profile of asterotocin mRNA in *A. rubens* 186

3.3.3.1. Radial nerve cord and circumoral nerve ring 187

3.3.3.2. Tube feet and marginal nerve 191

3.3.3.3. Digestive system 194

3.3.3.4. Body wall and pedicellariae 201

3.3.4. Asterotocin causes cardiac stomach eversion *in vivo* 203
3.3.5. Asterotocin causes cardiac stomach relaxation *in vitro* 207

3.4. Discussion 210
3.4.1. Discovery of asterotocin, a VP/OT-type peptide in *A. rubens* 210
3.4.2. Cardiac stomach eversion and retraction in *A. rubens* 211
3.4.3. Asterotocin: a regulator of cardiac stomach eversion 212
3.4.4. VP/OT-type neuropeptides: an evolutionarily ancient role in feeding behaviour? 216
3.4.5. VP/OT-type neuropeptides: an evolutionarily ancient role in osmoregulation and reproduction? 217
3.4.6. Asterotocin: a VP/OT-type peptide that causes muscle relaxation 218

4. Discovery and functional characterisation of a novel neurophysin-associated neuropeptide ("NGFFYamide") in the starfish *A. rubens*

4.1. Introduction 220

4.1.1. The NG peptide family 221

4.1.2. Functional roles of the NG peptides 224

4.1.3. Investigating the NG peptides in *A. rubens* 225

4.1.4. Aims and objectives 226

4.2. Methods 227

4.2.1. Bioinformatic identification of the NGFFYamide precursor in *A. rubens* 227

4.2.2. Cloning and sequencing of NGFFYamide precursor 227

4.2.3. Mass spectrometric confirmation of NGFFYamide 228

4.2.4. Localisation of the NGFFYamide precursor in *A. rubens* using mRNA *in situ* hybridisation (mRNA ISH) 229

4.2.5. Analysis of the *in vitro* effects of NGFFYamide on cardiac stomach preparations from *A. rubens* 230

4.2.6. Analysis of the *in vivo* effects of NGFFYamide in *A. rubens* 231

4.3. Results 232

4.3.1. Bioinformatic identification and cloning of the NGFFYamide precursor in *A. rubens* 232

4.3.2. Mass spectrometric confirmation of NGFFYamide 234

4.3.3. Expression profile of NGFFYamide mRNA in *A. rubens* 236

4.3.3.1. Radial nerve cord and circumoral nerve ring 236

4.3.3.2. Tube feet and marginal nerve 240

4.3.3.3. Digestive system 243

4.3.3.4. Apical muscle and coelomic epithelium 248

4.3.4. NGFFYamide causes cardiac stomach contraction *in vitro* 250

4.3.5. NGFFYamide causes cardiac stomach retraction *in vivo* 253

4.4. Discussion 256

4.4.1. Discovery of NGFFYamide, an NG peptide in *A. rubens* 256

4.4.2. NGFFYamide: a regulator of cardiac stomach contraction 257

4.4.3. The NG peptides: a conserved role in muscle myoactivity? 260

4.4.4. The NG peptides: neurophysin-associated neuropeptides 262

5. Discovery of the NGFFFamide receptor in the sea urchin *S. purpuratus* unites a bilaterian neuropeptide family

5.1. Introduction 263

5.1.1. Neuropeptide-S (NPS) 263

5.1.2. Crustacean cardioactive peptide (CCAP) 264

5.1.3. The NG peptides 265

5.1.4. NPS-type peptides, NG peptides and CCAP-type peptides: a bilaterian neuropeptide family? 267

5.1.5. Aims and objectives 269

5.2. Methods 270

5.2.1. Bioinformatic identification of an NPS/CCAP-type receptor in the sea urchin *S. purpuratus* 270

5.2.2. Cloning and sequencing of an NPS/CCAP-type receptor in the sea urchin *S. purpuratus* 270

5.2.3. Pharmacological characterisation of heterologously expressed NPS/CCAP-type receptor in the sea urchin *S. purpuratus* 271

5.2.4. Bioinformatic identification of novel NG peptide-type precursors and candidate NG peptide receptors in the echinoderms 272

5.3. Results 275

5.3.1. Cloning and sequencing of an NPS/CCAP-type receptor in the sea urchin *S. purpuratus* 275

5.3.2. NGFFFamide is a potent ligand for an NPS/CCAP-type receptor in the sea urchin *S. purpuratus* 277

5.3.3. Sequences of novel NG peptide precursors and candidate receptors in the echinoderms 279

5.3.3.1. NG peptide signalling system in the Echinoidea 282

5.3.3.2. NG peptide signaling system in the Holothurians 282

5.3.3.3. NG peptide signaling system in the Asteroidea 283

5.3.3.4. NG peptide signalling system in the Ophiuroidea 283

5.3.3.5. NG peptide signalling system in the Crinoidea 284

5.4. Discussion 290

5.4.1. An ancient duplication of a VP/OT-type neuropeptide signalling system gave rise to the NPS/NG/CCAP-type signalling systems 290

5.4.2. Characteristics of an ancestral neuropeptide precursor protein 295

5.4.3. An evolutionarily ancient role for the NPS/NG/CCAP-type neuropeptide signalling system? 298

5.4.4. Future directions 299

6. General discussion 300

6.1. *A. rubens*: providing novel insights into neuropeptide evolution 300
6.2. *A. rubens*: providing novel insights into neuropeptide function 301
6.3. The echinoderms: providing novel insights on neuropeptide evolution 302
6.4. Future directions: a platform for neuropeptide research 303

7. Acknowledgements 306

7.1. Identification of transcripts encoding neuropeptide precursors and neuropeptide receptors in the starfish *A. rubens* 306
7.2. Discovery and functional characterisation of a novel vasopressin/oxytocin (VP/OT)-type neuropeptide ("asterotocin") in the starfish *A. rubens* 307
7.3. Discovery and functional characterisation of a novel neurophysin-associated neuropeptide ("NGFFYamide") in the starfish *A. rubens* 307
7.4. Discovery of the NGFFFamide receptor in the sea urchin *S. purpuratus* unites a bilaterian neuropeptide family 308

8. References 309

9. Appendices 347

Tables and figures

Tables:

Table 1. Neuropeptide and polypeptide hormone precursors identified in *S. purpuratus* and *A. japonicus* 45
Table 2. ClustalX v.2.1 default colour scheme 64
Table 3. Summary of key peptide/precursor characteristics of bilaterian neuropeptide precursor families 76
Table 4. Summary of key peptide/precursor characteristics of neuropeptides/peptide hormones with multiple cysteine (C) residues 134
Table 5. Summary of key peptide/precursor characteristics of non-bilaterian neuropeptide precursor families 152
Table 6. NG peptides and candidate NPS/CCAP-type receptors previously reported in invertebrate deuterostomes 281

Figures:

Fig. 1.1. Processing and modification of a neuropeptide precursor 33
Fig. 1.2. Neuropeptide biosynthesis, processing and storage 35
Fig. 1.3. The Bilateria 38
Fig. 1.4. The echinoderms 42
Fig. 1.5. Diagram of a dissected starfish displaying the main anatomical features, with a particular focus on the digestive system 49
Fig. 2.1. *A. rubens* SALMFamide precursors 69
Fig. 2.2. The SALMFamide neuropeptide family 71
Fig. 2.3. *A. rubens* asterotocin precursor 78
Fig. 2.4. *A. rubens* NGFFYamide precursor 79
Fig. 2.5. *A. rubens* GnRH-type precursor 1 (ArGnRHP1) 83
Fig. 2.6. *A. rubens* GnRH-type precursor 2 (ArGnRHP2) 85
Fig. 2.7. *A. rubens* TRH-like precursor (ArTRHLPP) 89
Fig. 2.8. *A. rubens* TK-like precursor (ArTKLP) 93
Fig. 2.9. *A. rubens* CCK-like precursor (ArCCKLP) 98
Fig. 2.10. *A. rubens* OX-like precursors 103
Fig. 2.11. *A. rubens* LQ-like precursor (ArLQP) 107

Fig. 2.12. *A. rubens* KP-like precursor (ArKPLP) 111
Fig. 2.13. *A. rubens* SMS-like precursor (ArSMSLP) 115
Fig. 2.14. *A. rubens* MCH-like precursor (ArMCHLP) 118
Fig. 2.15. *A. rubens* CT-like precursor (ArCTLPP) 122
Fig. 2.16. *A. rubens* CRH-like precursor (ArCRHLP) 126
Fig. 2.17. *A. rubens* PDF-like precursor (ArPDFLP) 129
Fig. 2.18. *A. rubens* partial (3') pedal PP/orcokinin-like precursor (ArPPLNP1) 133
Fig. 2.19. *A. rubens* GPA2-type and GPB5-type precursors 138
Fig. 2.20. *A. rubens* bursicon-type precursors 140
Fig. 2.21. Conservation of bursicon-like, GPA2-type and GPB5-type subunits in representative phyla and species across the animal kingdom 141
Fig. 2.22. *A. rubens* GSS-like precursor (ArGSSLP) 143
Fig. 2.23. *A. rubens* relaxin-like precursor (ArRLP) 146
Fig. 2.24. *A. rubens* IGF-like precursors 150
Fig. 2.25. Conservation of insulin/IGF/relaxin-like peptide hormones in representative phyla and species across the Bilateria 151
Fig. 2.26. *A. rubens* AN peptide-type precursor (ArANPP) 155
Fig. 2.27. *A. rubens* Arnp11 precursor 156
Fig. 2.28. *A. rubens* Arnp15a-b precursors 158
Fig. 2.29. *A. rubens* Arnp18 precursor 159
Fig. 3.1. Conservation of vasopressin/oxytocin (VP/OT)-type peptides in representative phyla and species across the Bilateria 165
Fig. 3.2. Cardiac stomach assay 180
Fig. 3.3. *A. rubens* asterotocin precursor cDNA sequence 182
Fig. 3.4. *A. rubens* VP/OT-type receptor 183
Fig. 3.5. Mass spectrometric confirmation that asterotocin is present in the radial nerve cord of *A. rubens* 185
Fig. 3.6. Diagram of a transverse section of an arm of *A. rubens* 186
Fig. 3.7. Control experiment for asterotocin mRNA expression in the radial nerve cord in *A. rubens* 188
Fig. 3.8. Asterotocin mRNA expression in the radial nerve cord, lateral branches of the radial nerve cord and circumoral nerve ring in *A. rubens* 189
Fig. 3.9. Asterotocin mRNA expression in the tube feet and marginal nerve in *A. rubens* 192
Fig. 3.10. Diagram of the anatomy of the gut of *A. rubens* 195

Fig. 3.11. Asterotocin mRNA expression in the cardiac stomach of *A. rubens* 197
Fig. 3.12. Asterotocin mRNA expression in regions of the digestive system in *A. rubens* 199
Fig. 3.13. Asterotocin mRNA expression in the body wall and pedicellariae in *A. rubens* 202
Fig. 3.14. Asterotocin triggers cardiac stomach eversion in *A. rubens* 204
Fig. 3.15. Representative example of postural changes in the starfish *A. rubens* following injection of synthetic asterotocin 206
Fig. 3.16. Asterotocin is a potent stimulator of cardiac stomach relaxation in *A. rubens* 208
Fig. 4.1. The NG peptide family 222
Fig. 4.2. *A. rubens* NGFFYamide precursor cDNA sequence 233
Fig. 4.3. Mass spectrometric confirmation that NGFFYamide is present in the radial nerve cord of *A. rubens* 235
Fig. 4.4. Control experiment for NGFFYamide mRNA expression in the radial nerve cord in *A. rubens* 237
Fig. 4.5. NGFFYamide mRNA expression in the radial nerve cord, lateral branches of the radial nerve cord and circumoral nerve ring in *A. rubens* 238
Fig. 4.6. NGFFYamide mRNA expression in the tube feet and marginal nerve in *A. rubens* 241
Fig. 4.7. NGFFYamide mRNA expression in the cardiac stomach of *A. rubens* 244
Fig. 4.8. NGFFYamide mRNA expression in regions of the digestive system in *A. rubens* 246
Fig. 4.9. NGFFYamide mRNA expression in the apical muscle and coelomic epithelium of *A. rubens* 249
Fig. 4.10. NGFFYamide is a potent stimulator cardiac stomach contraction in *A. rubens* 251
Fig. 4.11. NGFFYamide triggers cardiac stomach retraction in *A. rubens* 254
Fig. 5.1. Gene contiguity in the sea urchin *S. purpuratus* 268
Fig. 5.2. *S. purpuratus* NPS/CCAP-type receptor 276
Fig. 5.3. The *S. purpuratus* NPS/CCAP-type receptor is activated by the NG peptides NGFFFamide, NGFFYamide and NGIWYamide 278
Fig. 5.4. The NPS/NG peptide precursors and receptors in deuterostomes 285
Fig. 5.5. NG peptide precursors identified in echinoderm species 287
Fig. 5.6. NPS/CCAP-type receptors identified in echinoderm species 288

Fig. 5.7. Evolution of the VP/OT-type and the NPS/NG/CCAP-type neuropeptide signalling systems in the Bilateria 292
Fig. 5.8. Characteristics of a common ancestral neuropeptide precursor protein 296

Appendices:

Fig. A1. Phylogenetic analysis of secretin-type receptors 347
Fig. A2. Phylogenetic analysis of rhodopsin β-type receptors 348
Fig. A3. Phylogenetic analysis of rhodopsin γ-type receptors 349
Fig. A4. Secretin-type receptors identified in *A. rubens* 350
Fig. A5. Rhodopsin β-type receptors identified in *A. rubens* 352
Fig. A6. Rhodopsin γ-type receptors identified in *A. rubens* 356
Fig. A7. Glycoprotein hormone-type receptors 360

Supplementary figures (on CD):

Fig. S1. *A. rubens* L-type SALMFamide precursor 1
Fig. S2. *A. rubens* F-type SALMFamide precursor 2
Fig. S3. *A. rubens* VP/OT-type precursor ("asterotocin" precursor) 3
Fig. S4. *A. rubens* VP/OT-type receptor ("asterotocin" receptor) 4
Fig. S5. *A. rubens* NGFFYamide precursor 6
Fig. S6. *A. rubens* NPS/CCAP-type receptor ("NGFFYamide" receptor) 7
Fig. S7. *A. rubens* GnRH-type precursor 1 (ArGnRHP1) 9
Fig. S8. *A. rubens* GnRH-type precursor 2 (ArGnRHP2) 10
Fig. S9. *A. rubens* GnRH-like receptor 1 (Ar_GnRHLR1) 11
Fig. S10. *A. rubens* GnRH-like receptor 2 (Ar_GnRHLR2) 13
Fig. S11. *A. rubens* TRH-like precursor (ArTRHLPP) 14
Fig. S12. *A. rubens* TK-like precursor (ArTKLP) 15
Fig. S13. *A. rubens* TK-like receptor 1 (Ar_TKLR1) 16
Fig. S14. *A. rubens* partial (5') TK-like receptor 2 (Ar_TKLR2) 18
Fig. S15. *A. rubens* CCK-like precursor (ArCCKLP) 19
Fig. S16. *A. rubens* CCK-like receptor (Ar_CCKLR) 21
Fig. S17. *A. rubens* OX-like precursor 1 (ArOXLP1) 22
Fig. S18. *A. rubens* OX-like precursor 2 (ArOXLP2) 23
Fig. S19. *A. rubens* OX-like receptor 1 (Ar_OXLR1) 24
Fig. S20. *A. rubens* partial (5') OX-like receptor 2 (Ar_OXLR2) 26

Fig. S21. *A. rubens* OX-like receptor 3 (Ar_OXLR3) 27
Fig. S22. *A. rubens* LQ-like precursor (ArLQLP) 29
Fig. S23. *A. rubens* LQ-like receptor 1 (Ar_LQLR1) 30
Fig. S24. *A. rubens* LQ-like receptor 2 (Ar_LQLR2) 32
Fig. S25. *A. rubens* KP-like precursor (ArKPLP) 34
Fig. S26. *A. rubens* partial (3') KP-like receptor 1 (Ar_KPLR1) 35
Fig. S27. *A. rubens* KP-like receptor 2 (Ar_KPLR2) 36
Fig. S28. *A. rubens* partial (5') KP-like receptor 3 (Ar_KPLR3) 37
Fig. S29. *A. rubens* KP-like receptor 4 (Ar_KPLR4) 38
Fig. S30. *A. rubens* KP-like receptor 5 (Ar_KPLR5) 39
Fig. S31. *A. rubens* KP-like receptor 6 (Ar_KPLR6) 41
Fig. S32. *A. rubens* KP-like receptor 7 (Ar_KPLR7) 43
Fig. S33. *A. rubens* KP-like receptor 8 (Ar_KPLR8) 44
Fig. S34. *A. rubens* KP-like receptor 9 (Ar_KPLR9) 45
Fig. S35. *A. rubens* SMS-like precursor (ArSMSLP) 46
Fig. S36. *A. rubens* partial (3') SMS-like receptor 1 (Ar_SMSLR1) 47
Fig. S37. *A. rubens* partial (5') SMS-like receptor 2 (Ar_SMSLR2) 48
Fig. S38. *A. rubens* SMS-like receptor 3 (Ar_SMSLR3) 49
Fig. S39. *A. rubens* MCH-like precursor (ArMCHLP) 51
Fig. S40. *A. rubens* MCH-like receptor 1 (Ar_MCHLR1) 52
Fig. S41. *A. rubens* MCH-like receptor 2 (Ar_MCHLR2) 54
Fig. S42. *A. rubens* CT-like precursor (ArCTLPP) 56
Fig. S43. *A. rubens* partial (5') CT-like receptor (Ar_CTLR) 57
Fig. S44. *A. rubens* CRH-like precursor (ArCRHLP) 58
Fig. S45. *A. rubens* CRH-like receptor 1 (Ar_CRHLR1) 59
Fig. S46. *A. rubens* partial (5') CRH-like receptor 2 (Ar_CRHLR2) 61
Fig. S47. *A. rubens* PDF-like precursor (ArPDFLP) 62
Fig. S48. *A. rubens* PDF-like receptor 1 (Ar_PDFLR1) 63
Fig. S49. *A. rubens* PDF-like receptor 2 (Ar_PDFLR2) 64
Fig. S50. *A. rubens* PP/orcokinin-like precursor 1 (ArPPLNP1) 66
Fig. S51. *A. rubens* GPA2-type precursor 1 67
Fig. S52. *A. rubens* GPA2-type precursor 2 (ArGPA2-2) 68
Fig. S53. *A. rubens* GPB5-type precursor 1 (ArGPB5-1) 69
Fig. S54. *A. rubens* GPB5-type precursor 2 (ArGPB5-2) 70
Fig. S55. *A. rubens* GPB5-type precursor 3 (ArGPB5-3) 71

Fig. S56. *A. rubens* partial (5') GPA2/5-type receptor 72
Fig. S57. *A. rubens* bursicon α-like precursor (ArBALP) 74
Fig. S58. *A. rubens* bursicon β-like precursor (ArBALP) 75
Fig. S59. *A. rubens* partial (3') bursicon-like receptor 76
Fig. S60. *A. rubens* GSS-like precursor (ArGSSLP) 79
Fig. S61. *A. rubens* relaxin-like precursor (ArRLP) 81
Fig. S62. *A. rubens* IGF1-like precursor (ArIGF1) 82
Fig. S63. *A. rubens* IGF2-like precursor (ArIGF2) 83
Fig. S64. *A. rubens* AN-peptide-type precursor (ArANPP) 84
Fig. S65. *A. rubens* Arnp11 precursor 86
Fig. S66. *A. rubens* Arnp15a precursor 87
Fig. S67. *A. rubens* Arnp15b precursor 88
Fig. S68. *A. rubens* Arnp18 precursor 89
Fig. S69. *L. variegatus* NGFFFamide precursor 90
Fig. S70. *P. parvimensis* NGIWY/NGVWYamide precursor 91
Fig. S71. *P. miniata* NGFFYamide precursor 92
Fig. S72. *O. victoriae* NGFFF/NGFFYamide precursor 93
Fig. S73. *A. filiformis* partial (5') NGFFFamide precursor 94
Fig. S74. *L. variegatus* NPS/CCAP-type receptor 95
Fig. S75. *P. parvimensis* NPS/CCAP-type receptor 96
Fig. S76. *P. miniata* NPS/CCAP-type receptor 98
Fig. S77. *A. mediterranea* NPS/CCAP-type receptor 99
Fig. S78. *F. serratissima* NPS/CCAP-type receptor 100

Abbreviations

ab	Aboral cardiac stomach
A.cal	*Aplysia californica*
ACEP-1	*Achatina* cardioexcitatory peptide-1
ACh	Acetylcholine
ACTH	Adrenocorticotropic hormone
ad	Adhesive region of tube foot
AKH	Adipokinetic hormone
Aj	*Apostichopus japonicus*
A.jap	*Apostichopus japonicus*
Ajnp	*Apostichopus japonicus* neuropeptide
am	Ampulla
AM	Apical muscle
A.med	*Antedon mediterranea*
ANPP	AN peptide-type precursor
Ar	*Asterias rubens*
A.rub	*Asterias rubens*
AST	Allatostatin
AT	Allatotropin
AVPL	Arginine vasopressin-like peptide
BALP	Bursicon-α-like precursor
BBLP	Bursicon-β-like precursor
BCIP	5-bromo-4-chloro-3’-indolyphosphate p-toluidine salt
B.flo	*Branchiostoma floridae*
BLAST	Basic local alignment search tool
BNP	Basiepithelial nerve plexus
BNR	Basal nerve ring
BW	Body wall
Ca^{2+}	Calcium
Calc	Calcitonin
cAMP	Cyclic adenosine monophosphate
CCAP	Crustacean cardioactive peptide
CCK	Cholecystokinin
CCKLP	Cholecystokinin-like precursor

cDNA	Complementary deoxyribonucleic acid
CE	Coelomic epithelium
C.ele	*Caenorhabditis elegans*
CG	Choriogonadotropin
cGMP	Cyclic guanosine monophosphate
CGRP	Calcitonin-gene related peptide
CHO	Chinese hamster ovary
C.int	*Ciona intestinalis*
CNP	Coelomic nerve plexus
CO_2	Carbon dioxide
con	Connective tissue
CONR	Circumoral nerve ring
CORT	Cortistatin
CRH	Corticotropin-releasing hormone
CRHLP	Corticotropin-releasing hormone-like precursor
CRHR	Corticotropin-releasing hormone receptor
CRISPR	Clustered regularly interspaced short palindromic repeats
CS	Cardiac stomach
CT	Calcitonin
C.tel	*Capitella teleta*
CTLPP	Calcitonin-like precursor protein
Cu^{2+}	Copper
DCV	Dense core vesicle
DH	Diuretic hormone
DIG-NTP	Dioxygenin nucleotide triphosphate
D.mel	*Drosophila melanogaster*
DMEM	Dulbecco's modified Eagle's medium
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
D.pul	*Daphnia pulex*
DTT	Dithiothreitol
EC_{50}	Half maximal effective concentration
ec	Ectoneural nerve plexus
EDTA	Ethylenediaminetetraacetic acid
ELH	Egg-laying hormone

ep	Epidermal layer
ER	Endoplasmic reticulum
ESI	Electrospray ionisation
EST	Expressed sequence tag
e-value	Expect value
FA	Formaldehyde
FBS	Fetal bovine serum
FNDI	Familial neurohypophyseal diabetes insipidus
FSH	Follicle-stimulating hormone
Gast	Gastrin
GDP	Guanosine diphosphate
GIP	Gastric inhibitory peptide
GnRH	Gonadotropin-releasing hormone
GnRHP	Gonadotropin-releasing hormone precursor
GPA2	Glycoprotein α–2
GPB5	Glycoprotein β-5
GPCR	G-protein coupled receptor
GPCRHMM	G-protein coupled receptor hidden Markov model
GSS	Gonad-stimulating substance
GSSLP	Gonad-stimulating substance-like precursor
GTP	Guanosine triphosphate
HCl	Hydrochloric acid
HDMS	High-definition mass spectrometer
HF	High fidelity
HMM	Hidden Markov model
H.sap	*Homo sapiens*
hy	Hyponeural nerve plexus
ICC	Immunocytochemistry
IGF	Insulin-like growth factor
INSL	Insulin-like
IP_3	Inositol triphosphate
ISH	*In situ* hybridisation
JH	Juvenile hormone
KCl	Potassium chloride
KP	Kisspeptin

KPLP	Kisspeptin-like precursor
lat	Lateral branch of radial nerve cord
LC	Liquid chromatography
LDCV	Large dense core vesicle
L.gig	*Lottia gigantea*
LGR	Leucine-rich repeat containing G-protein coupled receptor
LH	Luteinising hormone
LN	Longitudinal nerve
LQ	Luqin
LQLP	Luqin-like precursor
LSCPR	*Lymnaea stagnalis* conopressin receptor
lu	Lumen
LUQ	Left upper quadrant
LyCEP	*Lymnaea* cardioexcitatory peptide
MCH	Melanin-concentrating hormone
MCHLP	Melanin-concentrating hormone-like precursor
MEGA	Molecular evolutionary genetics analysis
$MgCl_2$	Magnesium chloride
MIH	Moult-inhibiting hormone
MN	Marginal nerve
mRNA	Messenger ribonucleic acid
mRNA ISH	Messenger ribonucleic acid *in situ* hybridisation
MS	Mass spectrometry
MSH	Melanocyte-stimulating hormone
mu	Mucosa
m/z	Mass/charge ratio
NaCl	Sodium chloride
NBT	Nitro-blue tetrazolium chloride
Ni^{2+}	Nickel
NKA	Neurokinin A
NKB	Neurokinin B
NJ	Neighbour joining
NMU	Neuromedin
NO	Nitric oxide
NOS	Nitric oxide synthase

NPF	Neuropeptide-F
NPS	Neuropeptide-S
NPSR	Neuropeptide-S receptor
NPY	Neuropeptide-Y
ODQ	1H-[1,2,4]oxadiazolo[4,3-a]quinoxalin-1-one
or	Oral cardiac stomach
ORF	Open reading frame
os	Ossicle
OT	Oxytocin
O.vic	*Ophionotus victoriae*
OX	Orexin
PAM	Point accepted mutation
pap	Papulae
PBS	Phosphate-buffered saline
PC	Pyloric caecae
PCR	Polymerase chain reaction
PD	Pyloric duct
PDF	Pigment-dispersing factor
PDFLP	Pigment-dispersing factor-like precursor
PDH	Pigment-dispersing hormone
ped	Pedicellariae
PFA	Paraformaldehyde
PK	Pyrokinin
PLC	Phospholipase-C
PP	Pedal peptide
PPLNP	Pedal peptide-like neuropeptide precursor
PS	Pyloric stomach
PTH	Parathyroid hormone
qPCR	Quantitative polymerase chain reaction
RC	Rectal caecae
Rel	Relaxin
RIA	Radioimmunoassay
RLM	Radial longitudinal muscle
RLP	Relaxin-like precursor
RNA	Ribonucleic acid

RNase	Ribonuclease
RNC	Radial nerve cord
RPC	Radial perihaemal canal
RSP	Regulated secretory pathway
RTK	Tyrosine kinase receptor
S1	SALMFamide 1
S2	SALMFamide 2
s.e.m	Standard error of the mean
SK	Sulfakinin
S.kow	*Saccoglossus kowalevskii*
SMS	Somatostatin
SMSLP	Somatostatin-like precursor
SNP	Sub-epithelial nerve plexus
SOAP	Short oligonucleotide analysis package
SOG	Somatostatin, opioid and galanin
SOP	*Styela* oxytocin-related peptide
Sp	*Strongylocentrotus purpuratus*
Sp-Ccap/NpSLr	*Strongylocentrotus purpuratus* crustacean cardioactive peptide/neuropeptide-S-like receptor
Sp-Echtcnr	*Strongylocentrotus purpuratus* vasopressin/oxytocin-type ("echinotocin") receptor
Spnp	*Strongylocentrotus purpuratus* neuropeptide
S.pur	Strongylocentrotus purpuratus
SSC	Saline-sodium citrate
SubP	Substance P
T_0	Time 0
$T_3/_4$	Thyroid hormones
TALEN	Transcription activator-like effector nuclease
tBLASTn	Translated nucleotide basic local alignment search tool
T.cas	*Tribolium castaneum*
TF	Tube feet
TK	Tachykinin
TRH	Thyrotropin-releasing hormone
TRHLP	Thyrotropin-releasing hormone-like precursor
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol

T.rub	*Takifugu rubripes*
TSH	Thyroid stimulating hormone
UPLC	Ultra performance liquid chromatography
VML	Visceral muscle layer
VNP	Visceral nerve plexus
VP	Vasopressin
VRR1	Vasopressin receptor-related receptor 1
WGD	Whole genome duplication
ZFN	Zinc finger nuclease

1. General introduction

1.1. Neuropeptides

Neuropeptides can be defined as short polypeptides (typically 5-50 amino acids in length) that are produced and released by neurons through the regulated secretory pathway (RSP) – these short polypeptides act on receptor proteins to regulate a diverse range of physiological processes and behaviours. Neuropeptides represent the largest class of signalling molecules in the nervous system; the mammalian genome encodes almost 70 "classic" neuropeptide precursors (Burbach, 2012). Thus, neuropeptides regulate an equally large and diverse range of physiological processes and behaviours, from the role of substance P in nociception (Zubrzycka and Janecka, 2000) to the role of oxytocin (OT) in reproduction (Gimpl and Fahrenholz, 2001). Neuropeptides can be pleiotropic; for example, neuropeptide-Y (NPY) has roles in regulating anxiety, feeding and in cardiovascular function (Thorsell and Heilig, 2002). However, the definition of a neuropeptide is not completely restrictive; neuropeptides can also be produced and released by non-neuronal cell types. One such example is mammalian somatostatin (SMS), which is also secreted by the gut (Pradayrol et al., 1978). Moreover, there are a number of other neuronal signalling molecules that are not considered classic neuropeptides but are considered neuropeptides in the broader perspective. These include peptide hormones, growth factors, granins and chemokines (Burbach, 2012).

1.1.1. Neuropeptide structural characteristics

Neuropeptides are derived from larger precursor proteins (or pre-prohormones) that are directly encoded in the genome. These larger precursor proteins are comprised of an N-terminal signal peptide and one or more bioactive neuropeptides that are usually flanked by dibasic/monobasic cleavage sites (**Fig. 1.1**) (Steiner, 1998). Bioactive neuropeptides may then undergo a number of post-translational modifications, which are often crucial for receptor binding and biological activity (Bennett et al., 1993).

The N-terminal signal peptide (typically 20-30 amino acids in length) has a tripartite structure consisting of a positively charged region (n-region), a hydrophobic core (h-region) and an uncharged but polar region (c-region) (von Heijne, 1985). The signal peptide is required for directing proteins into the lumen of the endoplasmic reticulum (ER) in the RSP (Rapoport et al., 1996; Zheng and Gierasch, 1996). Subsequently, the signal peptide is cleaved during passage from the ER membrane into the Golgi network at a consensus cleavage site in the c-region by an extracellular signal peptidase (von Heijne, 1985).

Bioactive neuropeptides are typically flanked by cleavage sites, which are recognised by pro-hormone convertase cleaving enzymes (Steiner, 1998). In vertebrates, cleavage most often occurs at dibasic cleavage sites comprised of the basic amino acids lysine (K) and arginine (R), whilst cleavage can also occur at monobasic cleavage sites (Devi, 1991; Rholam et al., 1995; Rouille et al., 1995; Seidah and Chretien, 1999). Importantly, cleavage site rules have also been proposed for insect neuropeptides, expanding our knowledge on neuropeptide prediction in non-vertebrate species (Veenstra, 2000).

Upon cleavage by pro-hormone convertases, neuropeptides are often subjected to post-translation modifications. The most common post-translational

modification is C-terminal amidation, which is often crucial for receptor binding and the biological activity of the secreted neuropeptide (Bennett et al., 1993). C-terminal amidation is catalysed by peptidyl-aminotransferase using a C-terminal glycine (G) residue as an amide donor for the preceding amino acid (Eipper et al., 1993). Other post-translational modifications contributing to biological activity include but are not limited to serine (S) phosphorylation (e.g. gastrin) (Varro and Dockray, 1993), N-linked acetylation (e.g. α-melanocyte-stimulating hormone (α-MSH)) (Dores et al., 1991), tyrosine (Y) sulfation (e.g. cholecystokinin (CCK)) (Mutt and Jorpes, 1968) and conversion of an N-terminal glutamine (Q) residue to a pyroglutamate (pQ) (e.g. thyrotropin-releasing hormone (TRH)) (Boler et al., 1969).

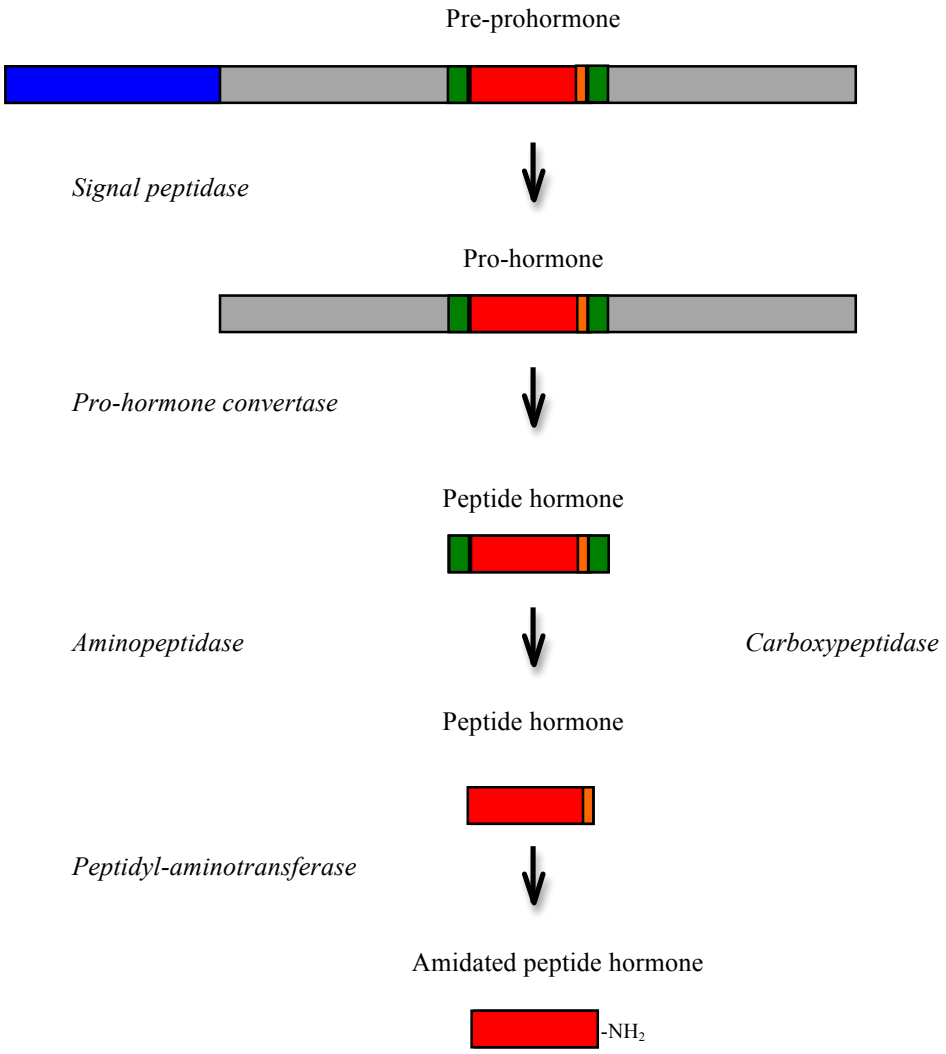


Fig. 1.1. Processing and modification of a neuropeptide precursor. A signal peptidase cleaves the signal peptide from a pre-prohormone to give rise to a pro-hormone. A pro-hormone convertase cleaves at dibasic/monobasic cleavage sites to give rise to a peptide hormone (neuropeptide). A carboxypeptidase and aminopeptidase cleave basic residues at the C-terminal and N-terminal respectively, giving rise to a bioactive peptide hormone. The peptide hormone can then undergo a number of potential post-translational modifications. For example, amidation is catalysed by peptidyl-aminotransferase, which uses a C-terminal glycine (G) residue as an amide donor for the preceding amino acid. The signal peptide is represented in blue, a neuropeptide is represented in red, a C-terminal glycine (G) residue is represented in orange and dibasic/monobasic cleavage sites are represented in green.

1.1.2. Neuropeptide release

Classic neuropeptides are produced and released by neurons via the RSP (Rapoport et al., 1996; Zheng and Gierasch, 1996). Neuropeptides are synthesised in the rough ER and are sorted in the trans-Golgi network into dense-core vesicles (DCVs) (**Fig. 1.2**) (Mains and Eipper, 1999). In DCVs, activation of proteolytic enzymes result in the processing of the neuropeptide precursor into one or more bioactive neuropeptides, which are released from the neuron via exocytosis at sites distant from the cell body (**Fig. 1.2**) (Brownstein et al., 1980; Lang et al., 1997; Tooze and Huttner, 1990).

However, in comparison to the release of neurotransmitters in small synaptic vesicles (Sudhof, 2012), far less is known about the release of neuropeptides. This is in part due to the small and slow nature of neuropeptide release, which makes electrophysiological studies more difficult (van den Pol, 2012). The release of neuropeptides from DCVs has been most extensively studied for the mammalian neurohypophyseal peptides vasopressin (VP) and OT, in which the presence of large DCVs (LDCVs) with high and measurable amounts of peptide has provided a suitable model system (van den Pol, 2012). In VP and OT release, the amount of neuropeptide released per spike increases with spike frequency (Dreifuss et al., 1971; Gainer et al., 1986). This has been shown to be due to an increase in cytoplasmic calcium (Ca^{2+}) in axon terminals, which is key to increasing the likelihood of LDCV exocytosis (Bondy et al., 1987; Jackson et al., 1991; Muschol and Salzberg, 2000).

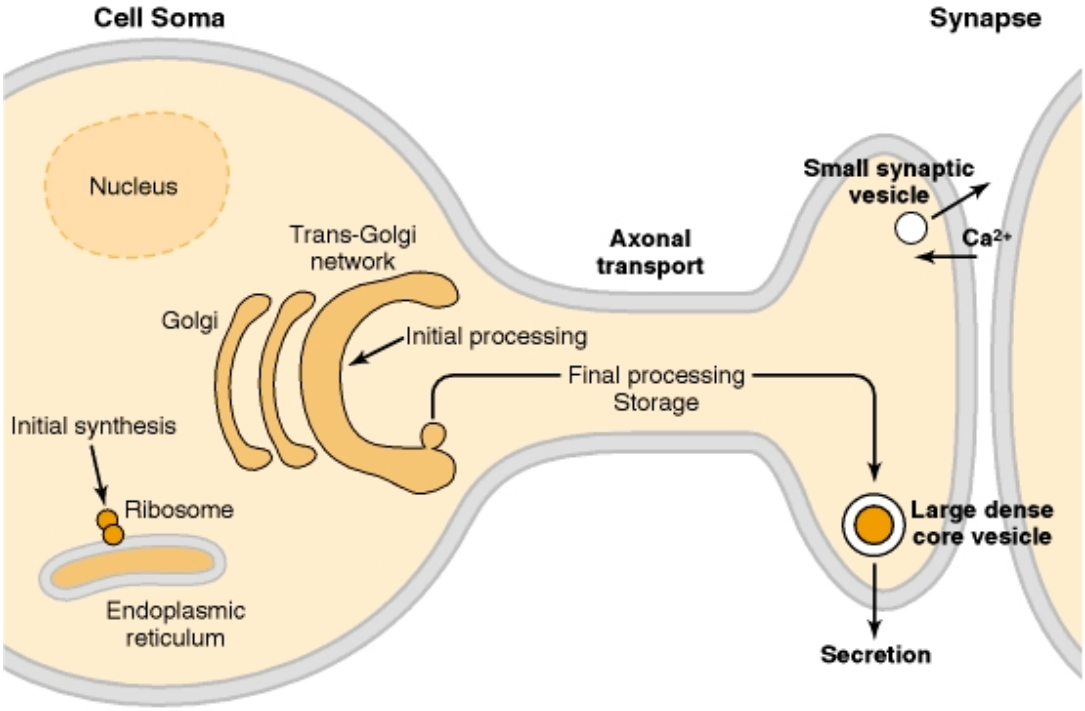


Fig. 1.2. Neuropeptide biosynthesis, processing and storage. Neuropeptide precursors are synthesised on ribosomes in the rough endoplasmic reticulum (ER) and processed through the Golgi apparatus. Neuropeptides are packaged into dense core vesicles (DCVs) or large dense core vesicles (LDCVs), which are axonally transported to the site of release. Neuropeptides are then released through the process of exocytosis. Figure from Mains and Eipper, 1999.

1.2. G-protein coupled receptors (GPCRs)

Neuropeptides, with a few notable exceptions including insulin-like (INSL) peptides that signal through tyrosine kinase receptors (RTKs), bind to and activate G-protein coupled receptors (GPCRs). GPCRs are a large and diverse family of membrane-bound receptors comprised of an extracellular N-terminus, seven-transmembrane spanning α-helices, an intracellular C-terminus and three inter-helical loops on each side of the membrane. GPCRs are classified into five families – rhodopsin, secretin, glutamate, adhesion and frizzled/taste-2 (Fredriksson et al., 2003). GPCRs bind a diverse range of ligands including photons, organic odorants, nucleotides, peptides, lipids and proteins (Bockaert and Pin, 1999). Furthermore, GPCRs represent a major target for drug discovery, with pharmaceuticals developed for a range of human diseases (Drews, 2000; Hopkins and Groom, 2002; Howard et al., 2001).

GPCRs function through intracellular heterotrimeric G-proteins, functioning as molecular switches linking ligand binding to intracellular signal transduction pathways. However, despite the large and diverse nature of the GPCR superfamily, there are relatively few G-proteins in comparison (Downes and Gautam, 1999). Heterotrimeric G-proteins are comprised of three subunits – an α-, β- and γ-subunit – divided into four main classes based on α-subunit specificity ($G\alpha_s$, $G\alpha_i$, $G\alpha_q$ and $G\alpha_{12}$) (Simon et al., 1991). Upon ligand binding, GPCRs undergo a conformational change to allow for G-protein binding and subsequent activation. The Gα-subunit cycles from an inactive guanosine diphosphate (GDP)-bound form to an active guanosine triphosphate (GTP)-bound form, which destabilises the Gα/GPCR complex to allow for the Gβ/Gγ dimer to interact with downstream effector proteins (Oldham and Hamm, 2008).

GPCRs activated by neuropeptides belong to three families/classes – the rhodopsin-β, rhodopsin-γ and secretin families (Fredriksson et al., 2003). The rhodopsin-β family in humans comprises a total of thirty-six receptors including the tachykinin (TK) receptors, CCK receptors, TRH receptors, gonadotropin-releasing hormone (GnRH) receptors and vasopressin/oxytocin (VP/OT)-type receptors (Fredriksson et al., 2003). The rhodopsin-γ family comprises three branches – the SOG (somatostatin, opioid and galanin-type) receptor cluster, the melanin-concentrating hormone (MCH) receptor cluster and the chemokine receptor cluster (Fredriksson et al., 2003). The secretin family comprises receptors including the calcitonin (CT) receptor, corticotropin-releasing hormone (CRH) receptors, the glucagon receptor, the gastric inhibitory polypeptide (GIP) receptor and the parathyroid hormone (PTH) receptors. The secretin family bind large peptide ligands and are characterised by an extracellular N-terminus between 60-80 amino acids in length comprising conserved disulphide bridges, which are crucial for ligand binding (Fredriksson et al., 2003).

1.3. Bilaterian neuropeptide signalling systems

1.3.1. The Bilateria

The Bilateria (animals with bilateral symmetry) are comprised of two super-phyla – the protostomes and the deuterostomes (**Fig. 1.3**). The protostomes (from Greek meaning “mouth first”) comprise the lophotrochozoans (e.g. molluscs and annelids) and the ecdysozoans (e.g. arthropods and nematodes). The deuterostomes (from Greek meaning “second mouth”) comprise the chordates (comprising vertebrates, cephalochordates and urochordates), the hemichordates (e.g. acorn worms), the echinoderms (e.g. starfish, sea urchins and sea cucumbers) and the xenacoelomorphs

(e.g. ciliated marine worms). The cnidarians, comprising hydrozoans (e.g. *Hydra*), cubozoans (e.g. comb jellyfish), scyphozoans (e.g. real jellyfish) and anthozoans (e.g. sea anemones and corals), are basal to the bilaterians in animal phylogeny (**Fig. 1.3**).

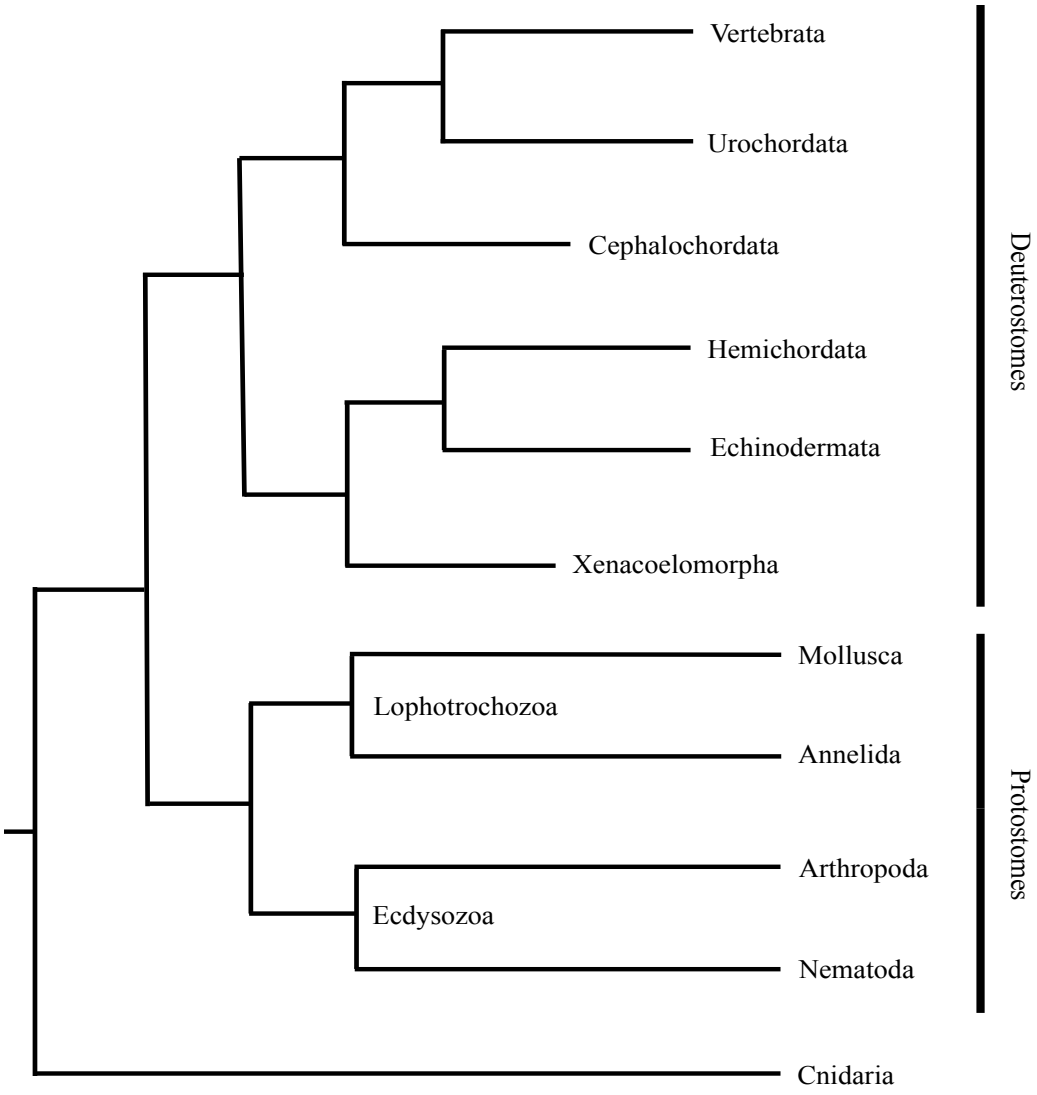


Fig. 1.3. The Bilateria. A simplified phylogenetic tree based upon the findings of Philippe et al., 2011 and Telford et al., 2006. The bilaterians are comprised of two super-phyla – the protostomes (lophotrochozoans and ecdysozoans) and the deuterostomes (vertebrates, urochordates, cephalochordates, hemichordates, echinoderms and xenacoelomorphs). The cnidarians are basal to the bilaterians.

1.3.2. The Bilateria: providing insights into the evolutionarily ancient origins of neuropeptide signalling systems

Neuropeptide signalling systems are key components of nervous systems in both the bilaterians (Jekely, 2013; Mirabeau and Joly, 2013) and the primitive cnidarians (Grimmelikhuijzen et al., 2004). Neuropeptides can be found in highly conserved families and can therefore be readily identified in evolutionary distant animals within the animal kingdom. One such example is the VP/OT-type neuropeptide signalling system; the mature peptide comprises a characteristic disulphide bridge between positions 1 and 6 crucial for bioactivity (Hruby et al., 1990; Sawyer, 1977) whilst the C-terminal region of the precursor contains a highly conserved neurophysin domain (comprising fourteen cysteine (C) residues) required for sorting in the RSP (de Bree, 2000; de Bree and Burbach, 1998).

However, such high sequence conservation is relatively uncommon and therefore establishing evolutionary relationships between neuropeptide signalling systems in evolutionarily distant phyla has historically proved difficult. Neuropeptides are highly diverse and are typically comprised of relatively short stretches of amino acids with relatively few conserved residues. Neuropeptides are also often repetitive, with the number and length of bioactive neuropeptides often changing through evolutionary time, providing difficulties for multiple alignment-based molecular phylogeny analyses (Jekely, 2013).

Recent advances in comparative genomics/transcriptomics has transformed our understanding of neuropeptide diversity and its functional significance across the animal kingdom. The increasing availability of genomic/transcriptomic sequence data has meant that neuropeptides from evolutionarily distant phyla have been shown to be orthologous and of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and

Joly, 2013). However, what is the importance of understanding whether, for example, an arthropod neuropeptide is related to a vertebrate neuropeptide?

Firstly, understanding the evolutionary relatedness of neuropeptides allows us to gain novel insights into neuropeptide function. For example, the sequencing of the genome of the fruit fly *Drosophila melanogaster* has led to the identification of a number of neuropeptide signalling systems (Brody and Cravchik, 2000; Hewes and Taghert, 2001; Vanden Broeck, 2001) that have provided important insights into the evolutionarily ancient roles of neuropeptide families. One such example is neuropeptide-F (NPF), which has been shown to regulate processes including foraging, feeding, alcohol sensitivity and aggression in *D. melanogaster* (Nässel and Wegener, 2011). Interestingly, it has been shown that NPF is related to NPY, which has been shown to regulate a range of processes including appetite and food consumption in vertebrates (Beck, 2006). It therefore appears that the NPF/NPY signalling system may have an evolutionarily ancient role in regulating feeding behaviours in the bilaterians.

Secondly, understanding the evolutionary relatedness of neuropeptides also allows for synergistic exchange of knowledge and resources between various fields of research. Key examples illustrating this have come from the genome of the nematode *Caenorhabditis elegans*; neuropeptides including adipokinetic hormone (AKH), pyrokinin (PK) and sulfakinin (SK) have been shown to be orthologous to vertebrate GnRH, neuromedin U (NMU) and CKK respectively through the identification of receptors that bind these neuropeptides (Janssen et al., 2008; Lindemans et al., 2011; Lindemans et al., 2009).

Finally, understanding the evolutionary relatedness of neuropeptides, both within and between different phyla, may provide pathways to pharmaceutical (Torres-Lugo and Peppas, 2000) and agrochemical (Bendena, 2010) discovery by informing development of synthetic non-peptidic receptor agonists and antagonists.

1.4. The echinoderms

The echinoderms (Phylum: Echinodermata) are marine animals, which, together with the hemichordates and the xenocoelomorphs, form a sister clade to the chordates (Bourlat et al., 2006; Bromham and Degnan, 1999; Philippe et al., 2011). The echinoderms comprise five extant classes – the echinoids (e.g. sea urchins), the holothurians (e.g. sea cucumbers), the asteroids (e.g. starfish), the ophiuroids (e.g. brittle stars) and the crinoids (e.g. feather stars and sea lilies) (**Fig. 1.4**.). The echinoids and holothurians form the echinozoan clade; the asteroids and ophiuroids form the asterozoan clade, whilst the crinoids are basal to the echinozoan and the asterozoan clades (O'Hara et al., 2014; Telford et al., 2014).

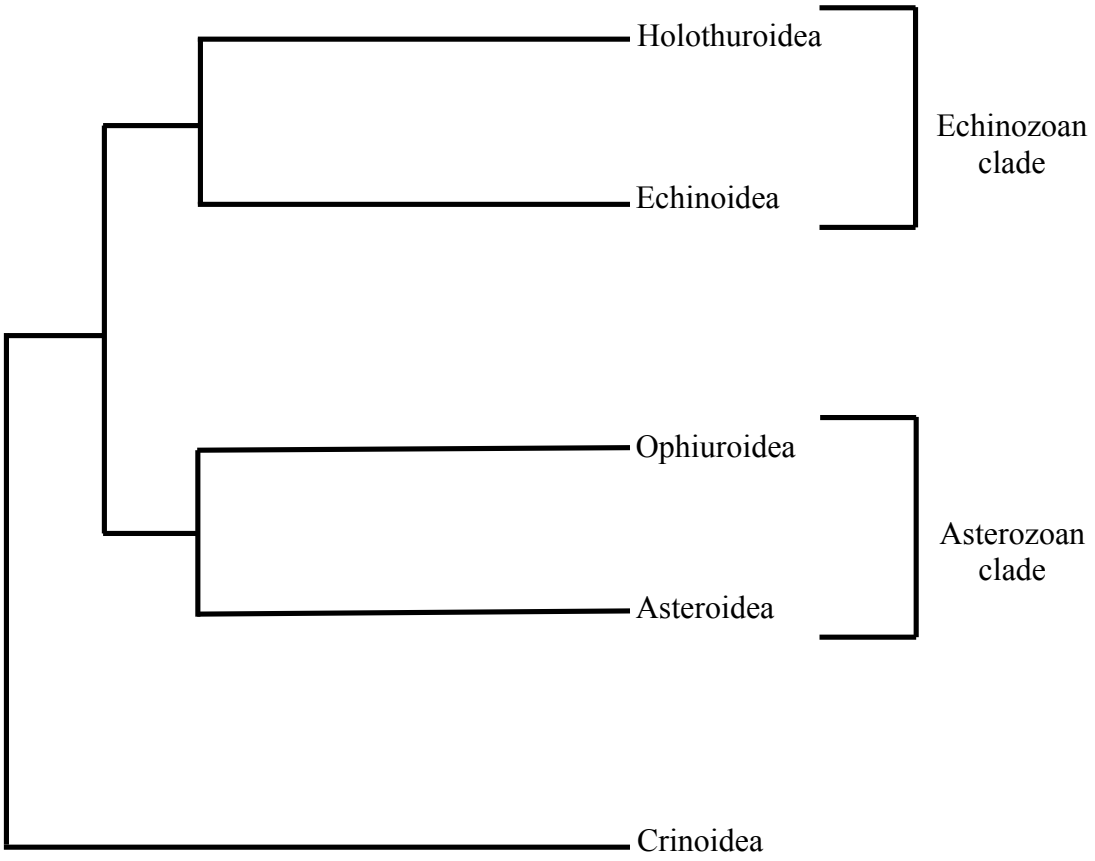


Fig. 1.4. The echinoderms. A simplified phylogenetic tree based upon the findings of O'Hara et al., 2014 and Telford et al., 2014. The holothurians (e.g. sea cucumbers) and echinoids (e.g. sea urchins) form the echinozoan clade. The ophiuroids (e.g. brittle stars) and asteroids (e.g. starfish) form the asterozoan clade. The crinoids (e.g. feather stars and sea lilies) are found basal to the echinozoan and asterozoan clades.

1.4.1. The echinoderms: a model system for neuropeptide research

The echinoderms provide a number of advantages as a model system for neuropeptide research over classic and long-standing model organisms including the fruit fly *D. melanogaster* and the nematode *C. elegans*, which continue to yield important insights into neuropeptide function in invertebrate species (Beets et al., 2012; Taghert and Nitabach, 2012). The echinoderms are of interest due to their phylogenetic position as a sister clade to the chordates; analysis of echinoderms therefore provides a basis for identifying synapomorphies shared within the deuterostome clade as well as factors that differentiate the echinoderms from the chordates. Moreover, the echinoderms provide a unique context for investigating the structure, function and evolution of neuropeptide signalling systems. The echinoderms possess a pentaradial morphological organisation along with the ability to autotomise, regenerate and rapidly and reversibly change the mechanical state of their body wall (Byrne, 2001; Patruno et al., 2001; Thorndyke et al., 2001; Wilkie, 2001; Wilkie, 2005).

Recent advances in comparative genomics/transcriptomics has transformed our knowledge and understanding of neuropeptide diversity across the animal kingdom. In 2006, the genome of the sea urchin *Strongylocentrotus purpuratus* was sequenced (Sodergren et al., 2006). The sea urchin genome project led to the identification of approximately 23,300 genes, with representatives of nearly all vertebrate gene families (Sodergren et al., 2006). Historically, the sea urchin has been an important model organism in developmental and systems biology. However, the sea urchin genome project has also allowed exploration into numerous regulatory networks including the defensome, adhesome, biomineralisation and the nervous system (Sodergren et al., 2006).

Analysis of the genome of the sea urchin *S. purpuratus* has resulted in the identification of a number of neuropeptide precursors (a SALMFamide precursor, a NGFFFamide precursor and a vasotocin-like precursor) and thirty-seven candidate neuropeptide receptors (Burke et al., 2006; Sodergren et al., 2006). Subsequently, analysis of 2026 expressed sequence tags (ESTs) from an *S. purpuratus* radial nerve cDNA library has led to the identification of twenty additional neuropeptide precursors (Rowe and Elphick, 2012), whilst a more recent analysis of sea urchin genome data has led to the identification of numerous additional neuropeptide precursors (Jekely, 2013; Mirabeau and Joly, 2013).

To date, investigation into the physiological functions of this repertoire of neuropeptides is limited. However, a VP/OT-type peptide ("echinotocin") has been shown to cause contraction of *in vitro* oesophagus and tube foot preparations from the sea urchin species *Echinus esculentus* (Elphick and Rowe, 2009). Importantly, transcriptome sequence data has also been obtained for a species belonging to another class of echinoderm in the echinozoan clade – a holothurian – the sea cucumber *Apostichopus japonicus* (Du et al., 2012). This has led to the identification of seventeen neuropeptide precursors and a number of candidate receptors (Rowe et al., 2014). However, when this study commenced, there was no publicly available transcriptome sequence data available for an asteroid, ophiuroid or crinoid species – generation and analysis of sequence data from these echinoderm classes would provide an opportunity to generate new insights into the evolution and diversity of neuropeptide signalling systems across the echinoderms.

S. purpuratus	Ref(s)	*A. japonicus*	Ref(s)
-	-	Stichopin, GN-19, GLRFA	*a, h*
NGFFFamide	*e*	NGIWYamide	*a, h*
Echinotocin	*e*		-
F-type SALMFamide	*e*	F-type SALMFamide	*b, h*
Spnp1 (L-type SALMFamide)	*f*	Ajnp1 (L-type SALMFamide)	*i*
Spnp2 (GnRHP)	*g*	-	-
Spnp3 (TRHLP)	*g*	Ajnp3 (TRHLP)	*l*
Spnp4 (CTLPP)	*g*	Ajnp4 (CTLPP)	*l*
Spnp5 (ANPP)	*g*	Ajnp5 (ANPP)	*l*
Spnp6 (PPLNP1)	*g*	-	-
Spnp7 (PPLNP2)	*g*	Ajnp7 (PPLNP2)	*l*
Spnp9	*g*	Ajnp9	*l*
Spnp16	*g*	Ajnp16	*l*
Spnp8, Spnp10-15, Spnp17-20	*g*	-	-
SpLQLP	*j*	AjLQLP	*l*
SpCRHLP	*k*	AjCRHLP	*l*
SpGPA2	*l*	AjGPA2	*l*
SpBALP	*d, j*	AjBALP	*l*
SpBBLP	*d*	AjBBLP	*l*

Table 1. Neuropeptide and polypeptide hormone precursors identified in *S. purpuratus* and *A. japonicus*. *BALP*, bursicon-α like precursor; *BBLP*, bursicon-β like precursor; *CRHLP*, corticotropin-releasing hormone-like precursor; *CTLPP*, calcitonin-like precursor; *GnRHP*, gonadotropin-releasing hormone-type precursor; *GPA2*, glycoprotein α-2-type precursor; *LQLP*, luqin-like precursor; *PPLNP*, pedal peptide-like precursor; *TRHLP*, thyrotropin-releasing hormone-like precursor. *a,* Iwakoshi et al., 1995; *b,* Ohtani et al., 1999; *c,* Elphick and Thorndyke, 2005; *d,* Burke et al., 2006; *e,* Elphick and Rowe, 2009; *f*, Rowe and Elphick, 2010; *g,* Rowe and Elphick, 2012; *h,* Elphick, 2012; *i*, Elphick et al., 2013; *j,* Jekely, 2013; *k*, Mirabeau and Joly, 2013; *l,* Rowe et al., 2014. Table adapted from Rowe et al., 2014.

1.5. The common European starfish *A. rubens*

The common European starfish *Asterias rubens* (Class: Asteroidea) is found throughout the coasts of the UK and through much of coastal Europe. However, the species also occurs in deeper waters off the northern coast of North America. Starfish reproduce by the process of free-spawning with gametes released into the environment from the gonads, which are located in pairs in each arm of the animal. In *A. rubens*, gonadal growth occurs through the autumn and winter, with gonadal size and oocyte diameter peaking in early spring prior to spawning (Barker and Nichols, 1983). *A. rubens*, whilst differing in shape and morphology from many of the approximate 1500 living species of asteroid, possesses a common external and internal anatomy to other starfish species with few exceptions (Ruppert et al., 2004).

1.5.1. *A. rubens* external anatomy

Many starfish species possess a striking pentaradial body plan with five arms (or rays) emanating from a central disk region (**Fig. 1.5**). However, some starfish species may possess more than five arms emanating from the central disk (e.g. the common sun star *Crossaster papposus*) (Ruppert et al., 2004). The spiny upper surface of the starfish is called the aboral surface, whilst the bottom surface is called the oral surface. The aboral surface contains structures including the anus, a madreporite that controls entry of seawater into the water vascular system, small jaw-like structures called pedicellariae that prevent colonisation of microorganisms and papulae for respiratory function (Cobb, 1978). The oral surface contains the mouth of the starfish, with an ambulacral groove extending radially from the mouth down each arm (Ruppert et al., 2004). The ambulacrum contains two to four rows of tubular projections termed tube feet, which drive the processes of locomotion, attachment

and feeding (Hennebert, 2010). The external anatomy of the starfish is supported by a body wall comprising an epidermal layer, a dermal layer (in which calcite ossicles are embedded) forming the endoskeleton, two layers of longitudinal and circular muscle and an inner coelomic epithelium (O'Neill, 1989).

1.5.2. *A. rubens* internal anatomy

The internal anatomy of starfish is comprised of a number of structures that form internal systems including a decentralised nervous system, a water vascular system, a digestive system and a haemal system.

The decentralised nervous system is comprised of five radial nerve cords located in each arm that extend outwards along the ambulacra and connect to a circumoral nerve ring. The circumoral nerve ring is located in the central disk region surrounding the mouth and is attached to and innervates the peristomial epidermis (Moore and Thorndyke, 1993). The circumoral nerve ring and radial nerve cord comprise an ectoneural and a hyponeural region (Cobb, 1970). The ectoneural nerve plexus is both sensory and motor in function, whilst the hyponeural nerve plexus is motor in function (Cobb, 1987; Moore and Thorndyke, 1993; von Hehn, 1970).

The tube foot driven processes of locomotion, attachment and feeding in addition to processes including respiration are powered by a water vascular system (Nichols, 1966; Nichols, 1972). The calcareous madreporite allows the movement of seawater into a single axial canal termed the stone canal and through to the ring canal located around the mouth of the starfish, which contains Tiedemann's bodies with a coelomocyte-like function (Holm et al., 2008). Pentamerous radial canals extend into each arm with perpendicular lateral canals allowing the movement of seawater into the tube feet, with ampullae acting as valves to power the tube feet.

Perhaps one of the most recognisable internal systems of the starfish is the

digestive system. Feeding in starfish species including *A. rubens* is a remarkable process by which the cardiac stomach is everted through the mouth over the digestible parts of the prey (e.g. mussels) through the relaxation of intrinsic muscles and extrinsic stomach retractor strands (Anderson, 1954). The eversion of the cardiac stomach enables starfish to feed on relatively large prey, with prey tissue partially digested externally. The cardiac stomach is then retracted through the contraction of intrinsic muscles and extrinsic stomach retractor strands (Basch, 1956; Burnett and Anderson, 1955) allowing the transport of partially digested tissue into the pyloric stomach, which, via pyloric ducts, is connected to a pair of pyloric caecae in each arm. From the pyloric stomach, a short tubular intestine joins to a rectum lined with out-pockets termed rectal caecae (Ruppert et al., 2004). The rectum opens at an anus in the centre of the aboral central disk (Ruppert et al., 2004).

The digestive system is linked to a haemal system, to allow for the transport and storage of nutrients from the pyloric caecae (Cuenot, 1948; Ruppert et al., 2004). The haemal system is comprised of three radial sets of vessels each consisting of a haemal ring in the disc and five radial vessels (Ruppert et al., 2004). The three haemal rings are the hyponeural haemal ring, the gastric haemal ring and the genital haemal ring, each connected by an axial sinus located in the coelomic cavity (Beijnink and Voogt, 1986; Broertjes and Posthuma, 1978). The heart is located on top of the axial sinus, above the three haemal rings and to one side of the madreporite (Ruppert et al., 2004).

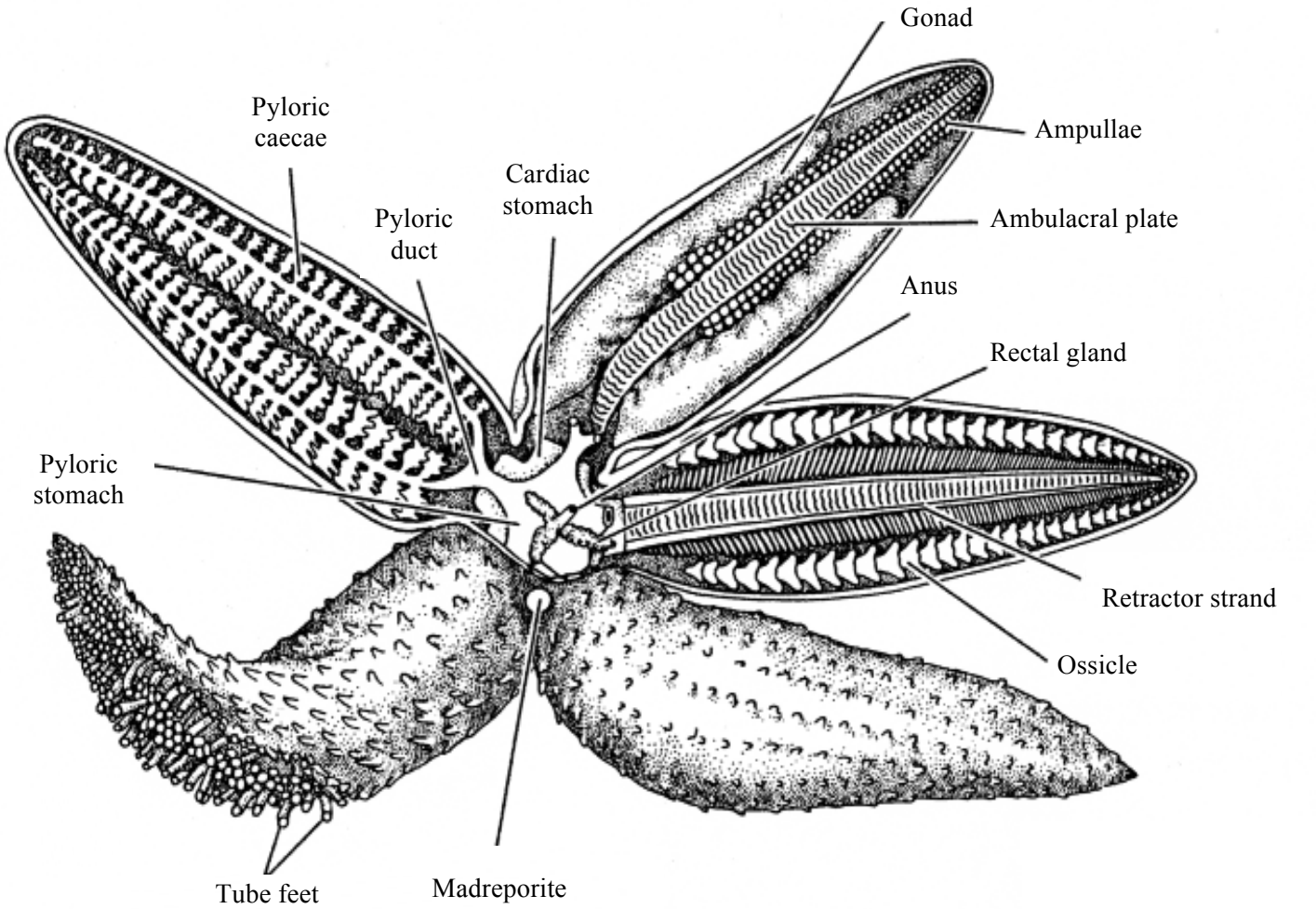

Gonad
Pyloric caecae
Ampullae
Cardiac stomach
Pyloric duct
Ambulacral plate
Anus
Rectal gland
Pyloric stomach
Retractor strand
Ossicle
Tube feet
Madreporite

Fig. 1.5. Diagram of a dissected starfish displaying the main anatomical features, with a particular focus on the digestive system. The digestive system comprises a cardiac stomach (with eversion and retraction controlled by retractor strands), a pyloric stomach, pyloric ducts, pyloric caecae, rectal caecae and an anus. The water vascular system comprises a madreporite, ampullae, ambulacral plate and tube feet. Also shown are ossicles (components of the body wall) and gonads (components of the reproductive system). Figure adapted from Brusca and Brusca, 2003.

1.5.3. *A. rubens*: a model system for neuropeptide research

The starfish *A. rubens* has been used as a model system for neuropeptide research for over twenty-five years – this was pioneered by the detection of FMRFamide-like immunoreactivity in the radial nerve and circumoral nerve ring in *A. rubens* (Elphick et al., 1989). This led to an effort to purify and identify peptides responsible for FMRF-like immunoreactivity in *A. rubens*, which resulted in the discovery of the first neuropeptides to be identified in an echinoderm species – the SALMFamides S1 (GFNSALMF-NH_2) and S2 (SGPYSFNSGLTF-NH_2) (Elphick et al., 1991a; Elphick et al., 1991b). Investigation into the distribution of S1 and S2 in *A. rubens* using immunocytochemistry (ICC) revealed expression in the nervous system and neuromuscular organs, including the digestive system, tube feet and apical muscle (Elphick et al., 1995; Newman et al., 1995a; Newman et al., 1995b). Subsequently, investigation into the pharmacological actions of S1 and S2 revealed that they cause relaxation of cardiac stomach, tube foot and apical muscle preparations *in vitro* (Elphick et al., 1995; Melarange and Elphick, 2003; Melarange et al., 1999) and cause cardiac stomach eversion *in vivo* (Elphick et al., 1995; Melarange et al., 1999).

In addition to providing pioneering insights into neuropeptide signalling systems in the echinoderms, *A. rubens* provides a number of further advantages as a model echinoderm system for neuropeptide research. Unlike other echinoderm species such as the sea urchin *S. purpuratus*, the starfish *A. rubens* is a common and easily obtained species in the UK and throughout much of coastal Europe, whilst *A. rubens* is also found in deeper waters off the northern coast of North America. Other closely related species of the genus *Asterias* also occur globally – *Asterias forbesi* is found along the Atlantic coast of the USA whilst *Asterias amurensis* is found along the coasts of Japan, China, Korea and Russia. Moreover, analysis of neuropeptide systems in *A. rubens* is important from an economical and environmental

perspective. Starfish species feed on economically important shellfish (e.g. mussels) harvested as foodstuffs (Agüera et al., 2012; Dare, 1982; Dolmer, 1998; Magnesen and Redmond, 2012). Importantly, the crown-of-thorns starfish *Acanthaster planci* feeds of reef building corals with periodic increases causing destruction of Pacific reef tracts (De'ath et al., 2012; Kayal et al., 2012; Timmers et al., 2012). Identification of neuropeptides that trigger cardiac stomach eversion (e.g. the SALMFamides S1 and S2) or retraction may therefore provide a basis for the development of non-peptidic antagonists that mimic or block the effects of endogenous neuropeptides, which could be used for chemical control of starfish species.

1.6. Aims and objectives

The aims and objectives of this thesis are represented in the proceeding results chapters:

(1) Identification of transcripts encoding neuropeptide precursors and neuropeptide receptors in the starfish *A. rubens*

The primary aim of the work reported in this this chapter is to sequence and assemble a neural transcriptome of the starfish *A. rubens* and utilise the transcriptome dataset to establish the species as a model system for neuropeptide research. This will provide a foundation to **(i)** investigate the evolutionary origins of neuropeptide signalling systems from the common ancestor of the Bilateria and **(ii)** investigate the physiological functions of neuropeptide signalling systems in **(a)** a starfish species, **(b)** an echinoderm species and even in **(c)** an ambulacrarian and **(d)** a non-chordate species, in many cases, for the very first time.

(2) Discovery and functional characterisation of a novel vasopressin/oxytocin (VP/OT)-type neuropeptide ("asterotocin") in the starfish *A. rubens*

The aim of the work reported in this chapter is to investigate the physiological function of a VP/OT-type peptide ("asterotocin") identified through the sequencing of a neural transcriptome of the starfish *A. rubens*. The sequence of the asterotocin precursor will be confirmed through cloning and sequencing, with the sequence of the asterotocin peptide and potential post-translational modifications confirmed through mass spectrometry. The physiological role(s) of asterotocin will be investigated utilising both *in vitro* and *in vivo* assays and the expression of asterotocin mRNA will be investigated through mRNA *in situ* hybridisation (ISH).

(3) Discovery and functional characterisation of a novel neurophysin-associated neuropeptide ("NGFFYamide") in the starfish *A. rubens*

The aim of the work reported in this chapter is to investigate the physiological function of an NG peptide ("NGFFYamide") identified through the sequencing of a neural transcriptome of the starfish *A. rubens*. The sequence of the NGFFYamide precursor will be confirmed through cloning and sequencing, with the sequence of the NGFFYamide peptide and potential post-translational modifications confirmed through mass spectrometry. The physiological role(s) of NGFFYamide will be investigated utilising both *in vitro* and *in vivo* assays and the expression of NGFFYamide mRNA will be investigated through mRNA ISH.

(4) Discovery of the NGFFFamide receptor in the sea urchin *S. purpuratus* unites a bilaterian neuropeptide family

---

The primary aim of the work reported in this chapter is to clone, sequence and pharmacologically characterise the receptor for the NG peptide NGFFFamide in the sea urchin *S. purpuratus*. The candidate NGFFFamide receptor is orthologous to neuropeptide-S (NPS)-type receptors in the tetrapod vertebrates and crustacean cardioactive peptide (CCAP)-type receptors in the protostomian invertebrates. Thus, characterisation of the NGFFFamide receptor will unify a bilaterian neuropeptide family that includes NPS-type peptides in vertebrates, NG peptides in deuterostomian invertebrates and CCAP-type peptides in protostomian invertebrates. The secondary aim of this chapter is to identify NG peptide precursors and candidate NPS/CCAP-type receptors utilising recently available genome/transcriptome data from a number of echinoderm species. In combination, these data will provide a foundation to pharmacologically characterise NG peptide receptors across the echinoderms.

2. Identification of transcripts encoding neuropeptide precursors and neuropeptide receptors in the starfish *A. rubens*

2.1. General introduction

Neuropeptides are evolutionarily ancient mediators of neuronal signalling that regulate a diverse range of physiological processes and behaviours across the animal kingdom. Previous phylogenetic studies aiming to understand the relationships of neuropeptide signalling systems outside of the vertebrates has been confined to analysing genome data from chordates including the urochordate *Ciona intestinalis* (Kamesh et al., 2008) and the cephalochordate *Branchiostoma floridae* (Nordstrom et al., 2008) and ecdysozoans including the fruit fly *Drosophila melanogaster* (Hewes and Taghert, 2001) and the nematode *Caenorhabditis elegans* (Nathoo et al., 2001).

However, recent advances in comparative genomics/transcriptomics has transformed our understanding of neuropeptide diversity across the animal kingdom, with a number of neuropeptide signalling systems having been recently traced back to the common ancestor of the Bilateria (Jekely, 2013; Mirabeau and Joly, 2013). The key to this approach has been the analysis of sequence data from a wider range of phyla across the animal kingdom; recent phylogenetic studies have utilised genome data from the urochordates (Dehal et al., 2002), cephalochordates (Putnam et al., 2008), hemichordates, echinoderms (Sodergren et al., 2006), lophotrochozoans (Simakov et al., 2013) and ecdysozoans (Adams et al., 2000; Colbourne et al., 2011; Richards et al., 2008) to provide important evidence for the evolutionarily ancient origins of a number of neuropeptide signalling systems (Jekely, 2013; Mirabeau and Joly, 2013).

2.1.1. The Ambulacraria

The echinoderms (e.g. starfish, sea urchins, sea cucumbers) and the hemichordates (e.g. acorn worms) form the ambulacrarians (Telford, 2006). The ambulacrarians are deuterostomian invertebrates and therefore occupy an intermediate position in animal phylogeny, bridging model protostomian invertebrates (e.g. *Drosophila* and *C. elegans*) to the vertebrates. Both *Drosophila* and *C. elegans* belong to the ecdysozoan clade and are therefore not representative of invertebrates as a whole. Therefore, the investigation of genome/transcriptome data available for the ambulacrarians has been crucial in providing novel insights into the evolutionarily ancient origins of neuropeptide signalling systems.

2.1.2. Ambulacrarian neuropeptide signalling systems: the hemichordates

The generation and analysis of genome data from the hemichordate *Saccoglossus kowalevskii* has recently proved invaluable in identifying neuropeptide signalling systems of a bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). For example, the identification of an orexin (OX)-type precursor in *S. kowalevskii* has unified a bilaterian neuropeptide family that is comprised of OX-type precursors in the deuterostomes and allatotropin (AT)-type precursors in the protostomes. OX's are hypothalamic neuropeptides that were first discovered as regulators of feeding behaviour (de Lecea et al., 1998; Sakurai et al., 1998) and have subsequently been implicated as key regulators in the processes of sleep and wakefulness in mammals (Sakurai, 2007). AT was first isolated through its regulatory function in stimulating the synthesis and secretion of juvenile hormone (JH) from the *corpora allata* and has subsequently been implicated in regulating cardioacceleration and the inhibition of active ion transport in the midgut of the tobacco hornworm *Manduca sexta* (Kataoka

et al., 1989; Veenstra et al., 1994). It had been shown that OX and AT share orthologous receptors (Vuerinckx et al., 2011; Yamanaka et al., 2008). However, homology could not be deduced based on neuropeptide sequences but were deduced on the presence of conserved domains in the neuropeptide precursor proteins (Mirabeau and Joly, 2013). In this regard, hemichordate genome data was utilised to identify conserved "cryptic" domains outside of peptide regions in the *S. kowalevskii* OX precursor and all protostomian AT precursors – this was not seen previously as the cryptic domain appears to have been lost in the chordates. This suggests that the OX and AT neuropeptide signalling systems are orthologous and can be traced back to the common ancestor of the Bilateria (Mirabeau and Joly, 2013).

2.1.3. Ambulacrarian neuropeptide signalling systems: the echinoderms

The generation and analysis of genome/transcriptome data from the sea urchin *S. purpuratus* (Burke et al., 2006; Rowe and Elphick, 2012) and transcriptome data from the sea cucumber *Apostichopus japonicus* (Rowe et al., 2014) have also recently proved invaluable in identifying neuropeptide signalling systems of a bilaterian origin. One such example is the identification of a thyrotropin-releasing hormone (TRH)-like precursor in *S. purpuratus*; this was the first TRH-like precursor to be identified in an invertebrate species and therefore allowed the neuropeptide signalling system to be traced back at least as far as the common ancestor of the deuterostomes (Rowe and Elphick, 2012).

The availability of genome/transcriptome data has also allowed for functional studies into the roles of novel neuropeptides in the echinoderms. For example, the identification of a vasopressin/oxytocin (VP/OT)-type peptide ("echinotocin") in *S. purpuratus* has shown that echinotocin causes contraction of *in vitro* tube foot and

oesophagus preparations from the sea urchin species *Echinus esculentus* (Elphick and Rowe, 2009).

In addition to sequence data for an echinoid species (*S. purpuratus*) and a holothurian species (*A. japonicus*), the genome of an asteroid – the starfish species *Patiria miniata* – has also been sequenced and has been made publically available (http://www.echinobase.org/Echinobase/pm). The utilisation of *P. miniata* genome data has, for example, provided interesting insights into the evolution and diversity of the SALMFamide family of neuropeptides (Elphick et al., 2013).

2.1.4. *A. rubens*: a model system for neuropeptide research

The importance of analysing and utilising genome/transcriptome data in the echinoderms has been highlighted by studies in the echinoid *S. purpuratus*, the holothurian *A. japonicus* and the asteroid *P. miniata*. However, the aim of this thesis is to utilise the common European starfish and asteroid *Asterias rubens* as a model system for neuropeptide research. *A. rubens* has a number of advantages over *P. miniata* as a model system for neuropeptide research. *A. rubens* has been utilised as a model system for over twenty-five years, with the SALMFamides S1 (GFNSALMF-NH_2) and S2 (SGPYSFNSGLTF-NH_2) the first neuropeptides to be identified in an echinoderm species (Elphick et al., 1991a; Elphick et al., 1991b); the identification of S1 and S2 has subsequently allowed for the investigation of the physiological function of S1 and S2 in this species (Elphick et al., 1995; Melarange and Elphick, 2003; Melarange et al., 1999). *A. rubens* is also a common and easily obtained species in the UK and throughout much of coastal Europe and deeper waters off the northern coast of North America, whilst *P. miniata* is found along the Pacific coast from Alaska to Mexico.

2.1.5. Aims and objectives

The aim of this chapter is to sequence and assemble a neural transcriptome of the starfish *A. rubens* to provide an opportunity to:

(1) Identify novel neuropeptide precursors and candidate receptors.

(2) Investigate the evolutionary origins of neuropeptide signalling systems from the common ancestor of the Bilateria.

(3) Provide an opportunity to investigate the physiological functions of neuropeptides in **(a)** a starfish species, **(b)** an echinoderm species and even in **(c)** an ambulacrarian and **(d)** a non-chordate species, in many cases, for the very first time.

2.2. Methods

2.2.1. Animal collection

Specimens of *A. rubens* were collected at low tide from the Thanet coast (Kent, UK) and transported to Queen Mary University of London (QMUL). Specimens of *A. rubens* were then maintained in a seawater aquarium at approximately 11 °C and fed with the bivalve mollusc *Mytilus edulis*.

2.2.2. RNA isolation

Radial nerve cords approximating 30 mg of tissue were dissected from a single male adult specimen and used for total RNA isolation using the SV Total RNA Isolation Kit (Promega). Total RNA was then transported to the Barts and the London Genome Centre in preparation for Illumina Hi-Seq.

2.2.3. Illumina Hi-Seq

Library preparation was performed using the TruSeq RNA Sample Preparation Kit v.2 (Illumina) at the Barts and the London Genome Centre. Transcriptome sequencing was performed on an Illumina HiSeq platform at the National Institute for Medical Research, with cBot used to generate clusters.

2.2.4. Transcriptome assembly

Raw sequence data was assembled using SOAPdenovo-Trans v.1.0 (http://soap.genomics.org.cn/SOAPdenovo-Trans.html), a short-read assembly method developed by the Beijing Genomics Institute (Li et al., 2008a). Contigs were assembled from reads with an overlap greater than 31 bp, which were then mapped back to the raw reads. The 326,816 contigs generated (with 16,316 over 1000 bp) were then set up for BLAST analysis using SequenceServer software (http://www.sequenceserver.com/), which is freely available to academic users (Priyam et al., in prep).

2.2.5. Bioinformatic identification of novel neuropeptide/peptide hormone precursors

To search for transcripts encoding putative neuropeptide or peptide hormone precursor proteins in the starfish *A. rubens*, a number of approaches were used:

(1) Sequences of neuropeptide or peptide hormone precursors identified in the sea urchin *S. purpuratus* (Burke et al., 2006; Elphick and Rowe, 2009; Elphick and Thorndyke, 2005; Jekely, 2013; Mirabeau and Joly, 2013; Rowe and Elphick, 2012), the sea cucumber *A. japonicus* (Rowe et al., 2014) and the starfish species *Asterina pectinifera* (Mita et al., 2009) were submitted individually as queries in tBLASTn searches of the contig database with the BLAST parameter expect (e)-value set to 1000. Contigs identified as encoding putative precursors were analysed after translation of their full length DNA sequence into a protein sequence using the ExPASy Translate tool (http://web.expasy.org/translate/). Proteins were assessed as potential precursors of secreted bioactive peptides by investigating: **(i)** the presence of a putative N-terminal signal peptide sequence using the SignalP 3.0 online server

(Bendtsen et al., 2004), **(ii)** the presence of putative dibasic/monobasic cleavage sites flanking putative bioactive peptide(s), with reference to known consensus cleavage motifs (Liu et al., 2006; Seidah and Chretien, 1999; Veenstra, 2000) and **(iii)** the presence, in some cases, of C-terminal glycine (G) residues that are putative substrates for amidation.

(2) A list of open reading frames (ORFs) generated from transcriptome assembly were analysed using a hidden Markov model (HMM), in which the scores used to rank precursor candidates are the expected density of cleavage inside a precursor (Mirabeau et al., 2007). The top 500 candidate sequences were then screened for the presence of an N-terminal signal peptide and short sequences flanked by dibasic cleavage sites and a C-terminal glycine (G) residue indicative of C-terminal amidation.

(3) A list of ORFs generated from transcriptome assembly were also analysed using the novel neuropeptide prediction tool NPsearch (https://rubygems.org/gems/NpSearch). NPsearch utilises the common characteristics of neuropeptides including the presence of an N-terminal signal peptide and short sequences flanked by dibasic cleavage sites and a C-terminal glycine (G) residue indicative of C-terminal amidation.

2.2.6. Multiple sequence alignments

Multiple sequence alignments of putative neuropeptides and/or selected regions of neuropeptide precursors were performed using ClustalX v.2.1 with a default colour scheme (**Table 2**) (Larkin et al., 2007). An asterisk denotes a single, fully conserved amino acid residue, a colon denotes strong conservation between groups of amino acid residues (≥ 0.5 Gonnet PAM 250 matrix) and a full stop denotes weak conservation between groups of amino acid residues (≤ 0.5 Gonnet PAM 250 matrix) (Gonnet et al., 1992). Multiple sequence alignments were performed using sequences from the starfish *A. rubens* and representative species from the vertebrates (e.g. *Homo sapiens* and *Takifugu rubripes*), the urochordates (e.g. *C. intestinalis*) the cephalochordates (e.g. *B. floridae*), the hemichordates (e.g. *S. kowalevskii*), echinoderms (e.g. *S. purpuratus*), the lophotrochozoans (e.g. *Capitella telata*, *Lottia gigantea* and *Aplysia californica*) and the ecdysozoans (e.g. *D. melanogaster* and *Daphnia pulex*) (Mirabeau and Joly, 2013).

Residue	Colour	Threshold/residue group
A, I, L, M, F, W, V	Blue	(+60 % WLVIMAFCHP)
R, K	Red	(+60 % KR) (+80 % KRQ)
N	Blue	(+50 % N) (+ 85 % NY)
C	Green	(+ 60 % WLVIMAFCHP)
C	Pink	(100 % C)
Q	Green	(+60 % KR) (+50 % QE) (+85 % QEKR)
E	Magenta	(+60 % KR) (+50 % QE) (+85 % EQD)
D	Magenta	(+60 % KR) (+85 % KRQ) (+50 % ED)
G	Orange	(0 % G)
H, Y	Cyan	(+60 % WLVIMAFCHP) (+85 % WYACPQ FHILMV)
P	Yellow	(+0 % P)
S, T	Green	(+60 % WLVIMAFCHP) (+50 % TS) (+85 % ST)

Table 2. ClustalX v.2.1 default colour scheme. In ClustalX v.2.1 multiple sequence alignments, amino acid residues are assigned a colour if the residue meets a specified criteria for the residue type. These specified criteria are represented as (+X %, xy) where X is the minimum percentage for a specified residue type represented as x or y.

2.2.7. Bioinformatic identification of candidate neuropeptide/peptide hormone receptors

To search for transcripts encoding candidate neuropeptide or peptide hormone receptors in the starfish *A. rubens*, sequences of neuropeptide or peptide hormone receptors identified in the sea urchin *S. purpuratus* (Burke et al., 2006; Jekely, 2013; Mirabeau and Joly, 2013) (or other representative bilaterian species if a receptor is absent in *S. purpuratus*) (Jekely, 2013; Mirabeau and Joly, 2013) were submitted individually as queries in tBLASTn searches of the contig database with the BLAST parameter e-value set to 1000. Contigs identified as encoding putative precursors were analysed after translation of their full length DNA sequence into protein sequence using the ExPASy Translate tool (http://web.expasy.org/translate/). The protein sequences were then analysed for the presence of seven transmembrane domains using GPCRHMM (http://gpcrhmm.sbc.su.se). Subsequently, the protein sequences were submitted individually in tBLASTn searches of The National Centre for Biotechnology Information (NCBI) databases (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine sequence similarity to known neuropeptide or peptide hormone receptors.

2.2.8. Neighbour joining (NJ) trees

Multiple sequence alignments of candidate *A. rubens* neuropeptide receptors with other bilaterian neuropeptide receptors were performed using ClustalW2 (Larkin et al., 2007). Neighbour joining (NJ) trees were constructed using MEGA v.6.06 (http://www.megasoftware.net) using bootstrapping (with 1000 bootstrap replications), a Poisson substitution model and uniform rates. Representative *H. sapiens* secretin-type neuropeptide receptors were used as an outgroup for rhodopsin β/γ-type NJ trees, whilst representative *H. sapiens* rhodopsin β-type neuropeptide receptors were used as an outgroup for secretin-type NJ trees (Mirabeau and Joly, 2013). NJ trees were annotated using the program FigTree v.1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

2.3. Results and Discussion

The generation and analysis of a neural transcriptome of the starfish *A. rubens* has led to the identification of thirty-five novel neuropeptide precursors and a number of candidate receptors (**Fig. A1-3**). These neuropeptide precursors are reported in four groups – precursors of neuropeptides that have previously been biochemically identified (see **chapter 2.3.1**), precursors of neuropeptides that belong to bilaterian neuropeptide families (see **chapter 2.3.2.**), precursors of neuropeptides and peptide hormones that comprise multiple cysteine (C) residues (see **chapter 2.3.3.**) and precursors of putative neuropeptides that do not share sequence similarity with any known bilaterian neuropeptide and peptide hormone families (see **chapter 2.3.4.**). These neuropeptide precursors, and potential candidate receptors, are discussed one-by-one (or in groups).

2.3.1. Precursors of neuropeptides that have been identified previously in *A. rubens* – the SALMFamides S1 and S2

The SALMFamides S1 (GFNSALMF-NH_2) and S2 (SGPYSFNSGLTF-NH_2) were the first neuropeptides to be identified in the echinoderms and were isolated from the starfish species *A. rubens* and *A. forbesi* (Elphick et al., 1991a; Elphick et al., 1991b). However, until now, the neuropeptide precursor(s) from which S1 and S2 are derived from in *A. rubens* were unknown.

The *A. rubens* L-type SALMFamide precursor is a 210-residue precursor protein comprising a predicted 23-residue N-terminal signal peptide and seven putative L-type SALMFamide-type peptides with the motif S/TxLxF/Y-NH_2 (where x is variable) bounded by dibasic cleavage sites (KR/RR) (**Fig. 2.1A; Fig. S1**). These include single copies of the L-type peptide sequences PAGASAFHSALSYG

(residues 67-80), AYHSALPFG (residues 94-102), AYHTGLPFG (residues 111-119), GFNSALMFG (or S1) (residues 136-144), LHSALPFG (residues 147-154), GYHSALPFG (residues 157-165) and GYHSGLPFG (residues 181-189) respectively. The presence of C-terminal glycine (G) residues on each of these peptide sequences is indicative of post-translational modifications giving rise to C-terminal amide groups on the mature peptides (e.g. PAGASAFHSALSY-NH_2).

The *A. rubens* F-type SALMFamide precursor is a 234-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide, seven putative F-type SALMFamide-type peptides with the motif F/YxF/Y-NH_2 (where x is variable) and one putative L-type SALMFamide-type peptide (or S2) bounded by dibasic cleavage sites (KR/RR) (**Fig. 2.1B; Fig. S2**). These include single copies of the F-type peptide sequences QQSDREREVEAAQTQFYPYG (residues 64-83), TDPRKASGGFTFG (residues 86-98), NIFGSYDFG (residues 161-169), AYGNNFSFG (residues 172-180), GMGVSSFSFG (residues 183-192), AFGDFSFG (residues 209-216), NNGLSSFTFG (residues 219-228) and one copy of the L-type peptide sequence SGPYSFNSGLTFG (or S2) (residues 141-153) respectively. The presence of C-terminal glycine (G) residues on each of these peptide sequences is indicative of post-translational modifications giving rise to C-terminal amide groups on the mature peptides (e.g. TDPRKASGGFTF-NH_2). It should also be noted that the presence of an N-terminal glutamine (Q) residue on the F-type polypeptide sequence QQSDREREVEAAQTQFYPY-NH_2 may be indicative of a post-translational modification, giving rise to an N-terminal pyro-glutaminyl (pQ) residue (pQQSDREREVEAAQTQFYPY-NH_2).

The discovery of the *A. rubens* L-type and F-type SALMFamide precursors has revealed that both S1 (GFNSALMF-NH_2) and S2 (SGPYSFNSGLTF-NH_2) are derived from different precursor proteins and provides further evidence that putative

L-type and F-type SALMFamide neuropeptides can also be derived from the same precursor protein (Elphick et al., 2013). In this regard, both SALMFamide precursors in *A. rubens* are similar to the orthologous SALMFamide precursors that have recently identified in the genome of the starfish species *P. miniata* (Elphick et al., 2013).

(A) L-type SALMFamide precursor

MKGQHLLAVAVVVVAGSFGIIEAYSPFGGYNRAPFDNVWVRADSMARGGSTGEDEAN
EQRMTGAKRPAGASAFHSALSYGKRGDDDSAEVERRAYHSALPFGKRTPIEKRAYHT
GLPFGKRDDEAAEQDAMMERRGFNSALMFGKRLHSALPFGKRGYHSALPFGKRLDTT
DEGDIIERRGYHSGLPFGKRATDDEAVNDILDQLRSEEN

(B) F-type SALMFamide precursor

MMVRFVALLGAVSLLVCQSAGLDAADVEEQDEFNKPYAPDSSYADLNALLGNNVPSL
HSASKRQQSDREREVEAAQTQFYPYGRRTDPRKASGGFTFGKRGQYFIPIPYEKREM
DEVNPYSVAKRDDELTGLEEYQASKRSGPYSFNSGLTFGKREPEKRNIFGSYDFGKR
AYGNNFSFGKRGMGVSSFSFGKRSGLEGEQMMPEDKRAFGDFSFGKRNNGLSSFTFG
KREGER

Fig. 2.1. *A. rubens* SALMFamide precursors. (A) L-type SALMFamide precursor. **(B)** F-type SALMFamide precursor. N-terminal signal peptides are represented in blue, putative SALMFamide neuropeptides are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green.

SALMFamide precursors have also been identified in other echinoderm classes (**Fig. 2.2**). The sea urchin *S. purpuratus* (Class: Echinoidea) has two precursors – an L-type SALMFamide precursor containing two putative L-type SALMFamides and an F-type SALMFamide precursor containing seven putative F-type SALMFamides (Elphick and Thorndyke, 2005; Rowe and Elphick, 2010). Similarly, the sea cucumber *A. japonicus* (Class: Holothuroidea) has two precursors – an L-type SALMFamide precursor containing three putative L-type SALMFamides and an F-type SALMFamide precursor containing five putative F-type SALMFamides and three putative L-type SALMFamides (Elphick et al., 2013).

However, until recently, it was not known whether the occurrence of an L-type and largely F-type SALMFamide precursor was an ancestral feature of the SALMFamide neuropeptide family. The generation and analysis of transcriptome data from the brittle star *Ophionotus victoriae* (Class: Ophiuroidea) and the feather star *Antedon mediterranea* (Class: Crinoidea) has shed light on this issue (**Fig. 2.2**) (Elphick et al., 2015). Similarly to the echinoids, holothurians and asteroids, the ophiuroid *O. victoriae* has two SALMFamide precursors – an L-type SALMFamide precursor containing two putative L-type SALMFamides and a putative F-type SALMFamide and an F-type SALMFamide precursor containing eleven putative F-type SALMFamides and a putative L-type SALMFamide. However, interestingly, in the crinoid *A. mediterranea*, only one SALMFamide precursor has been identified containing six putative L-type SALMFamides and six putative F-type SALMFamides. It has therefore been proposed that there was a single SALMFamide gene encoding both L-type and F-type SALMFamides before a duplication in a common ancestor of the echinozoans and asterozoans giving rise to an L-type SALMFamide precursor and an F-type SALMFamide precursor respectively (Elphick et al., 2015).

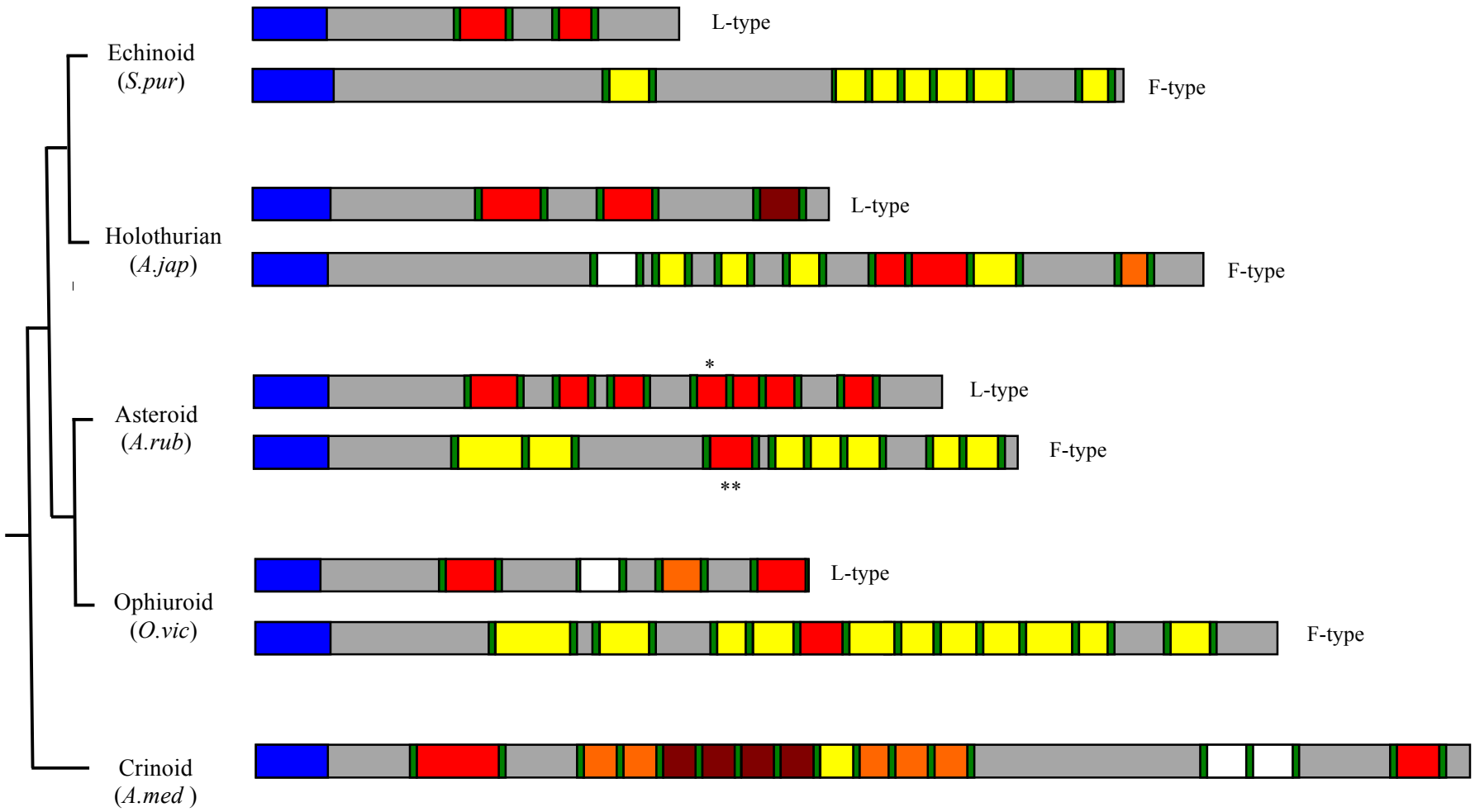

Echinoid
(S.pur)
L-type
F-type
Holothurian
(A.jap)
L-type
F-type
*
Asteroid
(A.rub)
L-type
F-type
**
Ophiuroid
(O.vic)
L-type
F-type
Crinoid
(A.med)

Fig. 2.2. The SALMFamide neuropeptide family. A phylogenetic schematic of the SALMFamide precursors across the echinoderms is shown. In an ancestral SALMFamide-type precursor in crinoids, the majority of the putative peptides have an FxLamide or an LxLamide-type motif and there is only one L-type SALMFamide and one F-type SALMFamide. However, as a consequence of specialisation following a predicted duplication of the ancestral-type precursor in the common ancestor of the echinozoans and asterozoans, two types of SALMFamide precursor have evolved – one that is predominantly comprised of L-type SALMFamides and another that is exclusively or predominantly comprised of F-type SALMFamides. Signal peptides are represented in blue. L-type SALMFamides with a C-terminal LxFamide motif or with an L-type-like motif (e.g. IxFamide) are represented in red. F-type SALMFamides with an FxFamide motif or with an F-type-like motif (e.g. YxFamide) are represented in yellow. SALMFamides with an FxLamide-type motif are represented in orange and SALMFamides with an LxLamide motif are represented in dark red. Peptides that do not conform to any of the four color-coded categories are represented in white. Dibasic or monobasic cleavage sites are represented in green. An asterisk denotes the position of S1 in the L-type precursor, whilst a double asterisk denotes the position of S2 in the F-type precursor in the starfish *A. rubens*. *A.jap = Apostichopus japonicus*; *A.med = Antedon mediterranea*; *A.rub = Asterias rubens*; *O.vic = Ophionotus victoriae*; *S.pur = Strongylocentrotus purpuratus*. Figure adapted from Elphick et al., 2015.

The physiological roles of the SALMFamide neuropeptides S1 and S2 in the starfish *A. rubens* have been previously investigated and have revealed a function in muscle relaxation both *in vitro* and *in vivo* (Elphick and Melarange, 2001; Elphick et al., 1995; Melarange and Elphick, 2003; Melarange and Elphick, 2003; Melarange et al., 1999). In this regard, S1 and S2 cause dose-dependent relaxation of cardiac stomach preparations *in vitro* (Melarange et al., 1999) whilst injection of S1 or S2 triggers cardiac stomach eversion *in vivo* (Elphick et al., 1995; Melarange et al., 1999).

Interestingly, S2 is approximately ten times more potent than S1 in cardiac stomach relaxation *in vitro* (Elphick and Melarange, 2001; Melarange and Elphick, 2003; Melarange and Elphick, 2003) whilst S2 is also more potent in inducing relaxation of tube feet and apical muscle preparations *in vitro* (Melarange and Elphick, 2003). It has been shown that the C-terminal regions of S1 and S2 are key to the differing potencies of both peptides (Otara et al., 2014). However, the discovery of the L-type and F-type SALMFamide precursors and fifteen putative SALMFamides has provided added complexity to previous functional studies. For example, it has recently been shown that the L-type SALMFamides AYHSALPF-NH_2 and GYHSGLPF-NH_2 bind copper (Cu^{2+}) and nickel (Ni^{2+}) and generate metal-linked dimers to form ternary complexes with LHSALPF-NH_2 (Jones et al., 2016). It has therefore been suggested that the formation of these ternary complexes may protect the SALMFamides from degradation from aminopeptidases during the process of extra-oral feeding in starfish (Jones et al., 2016). Thus, the discovery of a total of fifteen putative SALMFamides has provided new opportunities to investigate SALMFamide function and could provide a good model system for experimental analysis of the physiological significance of neuropeptide "cocktails" derived from the same precursor protein (Elphick et al., 2013).

To date, the relationship of echinoderm SALMFamide neuropeptides with neuropeptides that have been identified in other phyla remains unknown. Insights into orthologous neuropeptides in other phyla could be obtained through characterisation of the receptor(s) that mediate the effects of the SALMFamides.

2.3.2. Precursors of neuropeptides that belong to bilaterian neuropeptide families

Neuropeptide and peptide hormone precursors that have previously been identified as being of an ancestral bilaterian origin have been identified in the neural transcriptome of the starfish *A. rubens* (**Table 3**). In some cases, these novel neuropeptide and peptide hormone precursors have been identified in an echinoderm, an ambulacrarian and a non-chordate species for the very first time and will be discussed on a one-by-one basis (or in groups).

A. rubens precursor	Precursor family	Key peptide/precursor characteristics
Asterotocin	VP/OT-type	**Peptide:** Single peptide with conserved C residues (positions 1/6); C-terminal amidation **Precursor:** C-terminal neurophysin domain
NGFFYamide	NG peptide	**Peptides:** Multiple peptides with NG motif; C-terminal amidation **Precursor:** C-terminal neurophysin domain
ArGnRHP1-2	GnRH/AKH-type	Single peptide with N-terminal pQ residue; C-terminal amidated P motif
ArTRHLP	TRH-like	Multiple peptides with an N-terminal pQxP motif (x aromatic residue); C-terminal amidation
ArTKLP	TK-like	Single/multiple peptide(s) with amidated C-terminal pentapeptide motifs FxGLM-NH_2 (vertebrates) or FxGxR-NH_2 (invertebrates)
ArCCKLP	CCK/SK-like	Single/multiple peptide(s) with amidated C-terminal octapeptide motif DYMGWMDF-NH_2 (CCK) or heptapeptide motif DYGHMRF-NH_2 (SK)
ArOXLP1-2	OX/AT-like	Single peptide located after signal peptide with conserved C residues; C-terminal amidation
ArLQLP	LQ-like	**Peptide:** Single peptide with amidated C-terminal RF-NH_2 (lophotrochozoans) or RY-NH_2 motif (ecdysozoans) **Precursor:** C-terminal domain with two C residues separated by a 10-residue sequence
ArKPLP	KP-like	Single/multiple peptide(s) with amidated C-terminal LxF-NH_2 motif
ArSMSLP	SMS/AST-C-like	Single peptide with two conserved C residues located at N-terminus of precursor
ArMCHLP	MCH-like	Single peptide with two conserved C residues located at N-terminus of precursor
ArCTLPP	CT/DH31-like	Single peptide with two conserved C residues; C-terminal amidated P motif
ArCRHLP	CRH/DH44/ELH-like	Single/multiple peptide(s) with C-terminal amidated I/L motif
ArPDFLP	PDF-like	Single/multiple peptide(s) with C-terminal amidated V/A motif
ArPPLNP1	PP/orcokinin-like	Multiple peptides with SGFx motif (x hydrophobic residue)

Table 3. Summary of key peptide/precursor characteristics of bilaterian neuropeptide precursor families. *CCK/SK*, cholecystokinin/sulphakinin; *CRH/DH44/ELH*, corticotropin-releasing hormone/diuretic hormone-44/ egg-laying hormone; *CT/DH31*, calcitonin/diuretic hormone-31; *GnRH/AKH*, gonadotropin-releasing hormone/adipokinetic hormone; *KP*, kisspeptin; *LQ*, luqin; *MCH*, melanin-concentrating hormone; *OX/AT*, orexin/allatotropin; *PDF*, pigment-dispersing factor; *PP*, pedal peptide; *SMS/AST-C*, somatostatin/allatostatin-C; *TK*, tachykinin; *TRH*, thyrotropin-releasing hormone; *VP/OT*, vasopressin/oxytocin. Note that the table is not exhaustive and additional similarities and differences between neuropeptides/precursors are expanded on in the respective sub-headings.

2.3.2.1. VP/OT-type precursor ("asterotocin" precursor)

The *A. rubens* VP/OT-type precursor ("asterotocin" precursor) is a 147-residue precursor protein comprising a predicted 23-residue N-terminal signal peptide, one copy of the putative neuropeptide CLVQDCPEGG ("asterotocin") (residues 24-33) bounded by a dibasic cleavage site (KR) and a neurophysin domain (residues 44-121) containing a characteristic fourteen conserved cysteine (C) residues (**Fig. 2.3; Fig. S3**). The presence of a C-terminal glycine (G) residue is indicative of a post-translational modification, giving rise to a C-terminal amide group on the mature peptide (CLVQDCPEG-NH_2). It is noteworthy that asterotocin contains two cysteine (C) residues (residues 24 and 29), which are likely to form a disulphide bridge in accordance with VP/OT-type peptides in other phyla (de Bree and Burbach, 1998; Light and Du Vigneaud, 1958). Importantly, a candidate VP/OT-type receptor has also been identified in *A. rubens* (**Fig. 3.4; Fig. A5L; Fig. S4**). A detailed characterisation and discussion of the VP/OT-type neuropeptide signalling system is presented in **chapter 3.**

Vasopressin/oxytocin (VP/OT)-type precursor ("asterotocin" precursor)

MGMKSMVALWTGVLVALWVQSQACLVQDCPEGGKRSSYNTIRQCLSCGPGGLGQCVG
SAICCGNTFGCFLGTKETFVCREESQLSTPCEVVGETCESITDGKCVSNGFCCNERS
CSLDVACRETDTEQRDLKNRLKERLLDALLRQP

Fig. 2.3. *A. rubens* asterotocin precursor. An N-terminal signal peptide is represented in blue, the putative asterotocin neuropeptide (with cysteine (C) residues underlined) is represented red, a C-terminal glycine (G) residue is represented in orange and a dibasic cleavage site is represented in green. A neurophysin domain characteristic of VP/OT-type peptide precursors is represented purple, with a characteristic fourteen cysteine (C) residues underlined.

2.3.2.2. NGFFYamide precursor

The *A. rubens* NGFFYamide precursor is a 240-residue precursor protein comprising a predicted 23-residue N-terminal signal peptide, two copies of the putative neuropeptide NGFFYG ("NGFFYamide") (residues 124-130 and 132-137 respectively) bounded by dibasic cleavage sites (KR) and a neurophysin domain (residues 153-232) containing a characteristic fourteen cysteine (C) residues (**Fig. 2.4; Fig. S5**). The presence of C-terminal glycine (G) residues on each of these sequences is indicative of post-translational modifications giving rise to a C-terminal amide group on the mature peptide (NGFFY-NH_2). Importantly, a candidate NG peptide receptor has also been identified in *A. rubens* (**Fig. 5.6C; Fig. A5F; Fig. S6**). A detailed characterisation and discussion of the NG peptide signalling system is presented in **chapter 4**.

NGFFYamide precursor

MTMGSRSLLVTIVITVVIPSIWAGAIAGAQTQKIRRESRESGKYWPNSVGISDQQLR
QLLAHSLADSYSTSGASHIRGGDGDAGYIYDSRDQVDDTGTNEEEGERVIGSEVTSR
DSNPGTSKRNGFFYGKRNGFFYGKRSASTPGNANEVTQCIPCGPQNNGQCVMFGTCC
SYELGGCFFLTEEALPCVTSKSSSLCELSGLPCGDEGYGRCVADSVCCLPQEGSCHI
NAECGGKMTFQ

Fig. 2.4. *A. rubens* NGFFYamide precursor. An N-terminal signal peptide is represented in blue, putative NGFFYamide neuropeptides are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green. A neurophysin domain characteristic of NG peptide precursors is represented in purple, with a characteristic fourteen cysteine (C) residues underlined.

2.3.2.3. GnRH-type precursor 1 (ArGnRHP1)

The *A. rubens* gonadotropin-releasing hormone (GnRH)-type precursor 1 (ArGnRHP1) is a 121-residue precursor protein comprising a predicted 27-residue N-terminal signal peptide followed by a putative GnRH-type neuropeptide with the predicted peptide sequence QIHYKNPGWGPG (residues 28-40) termed ArGnRH followed by a dibasic cleavage site (KR) (**Fig. 2.5A; Fig. S7**). The presence of an N-terminal glutamine (Q) residue and a C-terminal glycine (G) residue are indicative of post-translational modifications giving rise to an N-terminal pyro-glutaminyl (pQ) residue and a C-terminal amide group on the mature peptide (pQIHYKNPGWGP-NH_2); this would be consistent with the occurrence of these post-translational modifications in GnRH-type peptides that have been identified in other species (Roch et al., 2011).

GnRH-type peptides are a family of neuropeptides with a widespread phylogenetic distribution indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). GnRH was first discovered in mammals as a reproductive regulator through its stimulatory effect on the release of the gonadotropins luteinising hormone (LH) and follicle-stimulating hormone (FSH) from the anterior pituitary gland (Kah et al., 2007). GnRH-type peptides have subsequently been identified in other vertebrates (Guilgur et al., 2006; Kavanaugh et al., 2008; Kim et al., 2011; Morgan and Millar, 2004), deuterostomian invertebrates including urochordates (Gorbman and Sower, 2003; Tsai, 2006), the cephalochordate *B. floridae* (Roch et al., 2014), the hemichordate *S. kowalevskii* (Mirabeau and Joly, 2013) and the echinoderm *S. purpuratus* (Rowe and Elphick, 2012). GnRH-type peptides have also been identified in the lophotrochozoans including a number of molluscs (Di Cristo et al., 2009; Iwakoshi et al., 2002; Tsai and Zhang, 2008; Zhang et al., 2000a) and annelids (Roch et al., 2011; Tsai and Zhang, 2008; Veenstra,

2011).

GnRH-type peptides have been found to be orthologous to adipokinetic hormone (AKH) and corazonin-type peptides identified in the ecdysozoans including a number of arthropods and nematodes (Lindemans et al., 2009; Lindemans et al., 2011). AKH was first discovered on account of its synthesis in the corpora cardiaca, which is considered the functional equivalent of the pituitary gland in vertebrates (O'Shea and Rayne, 1992). AKH has been shown to be involved in mobilising energy from the fat body during flight and locomotion in insects (Gade, 2004). More recently, however, AKH-type peptides have also been shown to have a central role in regulating fertility in the nematode *C. elegans*, which is indicative of GnRH/AKH-type peptides having an evolutionarily ancient role in regulating reproductive processes (Lindemans et al., 2009). However, whilst insect AKH has not yet been implicated with a role in reproductive functions, AKH-like expression has been detected in reproductive organs in various insect species including the female mosquito *Anopheles gambiae* (Kaufmann and Brown, 2006) and the fall armyworm *Spodoptera frugiperda* (Abdel-Latief and Hoffmann, 2007).

It is important to consider ArGnRHP1 in relation to GnRH-type precursors/peptides identified in other phyla (**Fig. 2.5B**). The presence of a single GnRH-type peptide located directly after the N-terminal signal peptide in ArGnRHP1 is a conserved feature of all GnRH-type precursors. Moreover, the presence of an N-terminal pyro-glutaminyl (pQ) residue and C-terminal amide group in the peptide is also a feature of all known GnRH-type peptides. However, there are a few exceptions, including AKH-type peptides in *C. elegans* (Lindemans et al., 2009) and insects (e.g. *S. frugiperda*) (Abdel-Latief and Hoffmann, 2007), which does contain a C-terminal amide group (Lindemans et al., 2011).

It is also important to consider ArGnRHP1 in relation to GnRH-type

precursors/peptides identified in other echinoderm species (**Fig. 2.5B**). ArGnRHP is only the second member of the GnRH-type neuropeptide family to be identified in an echinoderm after the identification of Spnp2 (SpGnRHP) in the sea urchin *S. purpuratus* (Roch et al., 2011; Rowe and Elphick, 2012). To date, a GnRH-type precursor has not been identified in a transcriptome dataset from the sea cucumber *A. japonicus* (Rowe et al., 2014). Interestingly, whilst GnRH-type peptides derived from ArGnRHP1 and SpGnRHP (pQVHHRFSGWRPG-NH_2) have all the characteristics of GnRH-type peptides identified in other phyla, SpGnRH also has an F (S/T) signature characteristic of protostomian GnRH-type peptides as opposed to chordate GnRH-type peptides (Lindemans et al., 2011).

To date, the physiological function of GnRH-type peptides in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of GnRH-type peptides. In this regard, candidate GnRH-type receptors have been identified in the sea urchin *S. purpuratus* (Burke et al., 2006). Importantly, however, two candidate GnRH-type receptors have also been identified in the starfish *A. rubens* and may facilitate investigation into the physiological functions of GnRH-type peptides in an echinoderm species (**Fig. A5B-C; Fig. S9-10**).

(A) Gonadotropin-releasing hormone (GnRH)-type precursor 1 (ArGnRHP1)

MADMRMLTLTSVLVSLLFMAEIQRCQGQIHYKNPGWGPGGKRSSHMTGSNVLRKRHW
RVESDQMGTDSMQKERNLIMLQEIAKSLAKQLVVPTSEDDTVLDQLTVDQWRQEADE
INDNGWN

(B) GnRH-type neuropeptide multiple sequence alignments

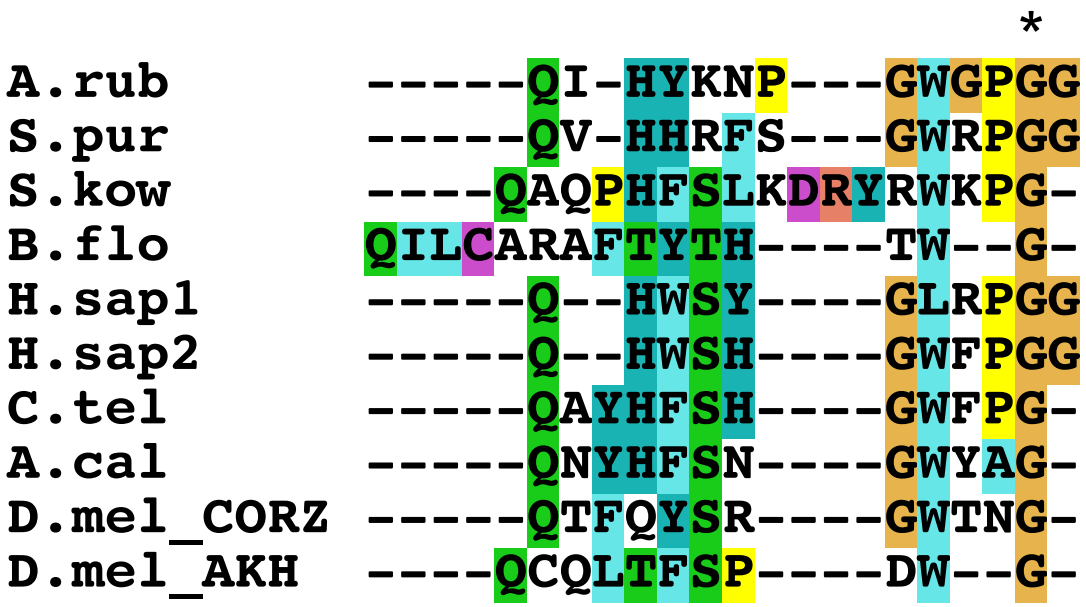


Fig. 2.5. ***A. rubens*** **GnRH-type precursor 1 (ArGnRHP1). (A)** Sequence of ArGnRHP1. An N-terminal signal peptide is represented in blue, a putative GnRH-type neuropeptide is represented in red, a C-terminal glycine (G) residue is represented in orange and a dibasic cleavage site is represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative GnRH-type neuropeptides derived from precursors of representative species across the Bilateria. *A.cal*, *Aplysia californica* GnRH-type precursor [GI: 325296898]; *A.rub*, *Asterias rubens* GnRH-type precursor; *B.flo*, *Branchiostoma floridae* GnRH-type precursor [GI: 568818760]; *C.tel*, *Capitella teleta* GnRH-type precursor [GI: 161294493]; *D.mel_AKH*, *Drosophila melanogaster* adipokinetic hormone (AKH) precursor [GI: 281365621]; *D.mel_CORZ*, *Drosophila melanogaster* corazonin (CORZ) precursor [GI: 386765761]; *H.sap1*, *Homo sapiens* GnRH1 [GI: 133908609]; *H.sap2*, *Homo sapiens* GnRH2 [GI: 109731929]; *S.kow*, *Saccoglossus kowalevskii* GnRH-type precursor [GI: 585702722]; *S.pur*, *Strongylocentrotus purpuratus* GnRH-type precursor [GI: 390361802].

2.3.2.4. GnRH-type precursor 2 (ArGnRHP2)

The *A. rubens* gonadotropin-releasing hormone (GnRH)-type precursor 2 (ArGnRHP2) is a 99-residue precursor protein comprising a predicted 23-residue N-terminal signal peptide and a putative GnRH-type peptide sequence (HNTFTMGGQNRWKAGG) termed ArGnRH2 followed by a dibasic cleavage site (KR) (**Fig. 2.6A; Fig. S8**). The presence of a C-terminal glycine (G) residue is indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide (HNTFTMGGQNRWKAG-NH_2). However, the absence of an N-terminal pyroglutamate (pQ) residue is not consistent with GnRH-type peptides that have been identified in other species (Roch et al., 2011).

ArGnRHP2 was identified through sequence similarity with Spnp12 in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012). However, upon the discovery of ArGnRH2 and alignments with ArGnRHP and Spnp2 (SpGnRHP), it is likely that Spnp12 is a GnRH-type neuropeptide precursor comprising a putative GnRH-type neuropeptide (HNTFSFKGRSRYFP-NH_2) (**Fig. 2.6B**). ArGnRHP2 and Spnp12 (SpGnRHP2) share superficial sequence similarity with GnRH-type peptides in the putative neuropeptide region comprising an F (S/T) signature characteristic of protostomian GnRH-type peptides as opposed to chordate GnRH-type peptides (Lindemans et al., 2011). Significantly, both ArGnRHP2 and Spnp12 (SpGnRHP2) have a highly conserved C-terminal domain, providing further evidence that the precursors are orthologous and belong to a family of GnRH-type neuropeptide precursors (**Fig. 2.6B**). The presence of two GnRH-type neuropeptides in both the starfish *A. rubens* and the sea urchin *S. purpuratus* add further complexity to the evolutionary origins of a neuropeptide signalling system that comprises GnRH-type, AKH-type and corazonin-type peptides. Accordingly, the gene duplications that gave rise to this diverse family of neuropeptides remains unclear (Roch et al., 2014).

As described previously, the physiological function of GnRH-type peptides in the echinoderms is currently unknown. However, the discovery of two GnRH-type peptides and two candidate GnRH-type receptors in the starfish *A. rubens* may facilitate investigation into the physiological functions of GnRH-type peptides in an echinoderm species (**Fig. A5B-C; Fig. S9-10**).

(A) Gonadotropin-releasing hormone (GnRH)-type precursor 2 (ArGnRHP2)

MGSYSVTATIYLALVLGSLVCSAHNTFTMGGQNRWKAGGKRSAPAGRPQQTFLDPSS
FSEDQQGETTITLREMLVDMRDYCSFLLKLLDNVRLPQTERK

(B) GnRH-type neuropeptide/precursor multiple sequence alignment

(i)

```
                    :   .    *    :  .*
A.rub_GnRH    QIHYKNPG---WGPGG
S.pur_GnRH    QVHHRFSG---WRPGG
A.rub_GnRH2   HNTFTMGGQNRWKAGG
S.pur_GnRH2   HNTFSFKGRSRYFPG-
```

(ii)

```
                :* :: :       *:          *: :      :
A.rub_GnRH    MLQEIAKSLAKQ-LVVPTSEDDTVLDQLT--VDQWRQEN-------
S.pur_GnRH    ILGDLRRAANRMRLLQLFNLSKTRLNDLNDATSNEPVYGDYLGTGL
A.rub_GnRH2   MLVDM-RDYCSF-LLK-------LLDNVR--LPQ------------
S.pur_GnRH2   MLTEL-RGYCDF-LLK-------LLDGVRPDLPQQRK---------
```

Fig. 2.6. *A. rubens* GnRH-type precursor 2 (ArGnRHP2). (A) Sequence of ArGnRHP2. An N-terminal signal peptide is represented in blue, a putative GnRH-type neuropeptide is represented in red, a C-terminal glycine (G) residue is represented in orange and a dibasic cleavage site is represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of **(i)** putative GnRH-type neuropeptides and **(ii)** a C-terminal domain of GnRH-type precursors in the starfish *A. rubens* and the sea urchin *S. purpuratus*. *A.rub_GnRH*, *Asterias rubens* GnRH-type precursor; *A.rub_GnRH2*, *Asterias rubens* GnRH-type precursor 2; *S.pur_GnRH*, *Strongylocentrotus purpuratus* GnRH-type precursor [GI: 390361802]; *S.pur_GnRH2*, *Strongylocentrotus purpuratus* GnRH-type precursor 2 [GI: 109403263].

2.3.2.5. TRH-like precursor (ArTRHLP)

The *A. rubens* thyrotropin-releasing hormone (TRH)-like precursor (ArTRHLP) is a 240-residue precursor protein comprising a predicted 23-residue N-terminal signal peptide and thirteen putative TRH-type peptides bounded by dibasic cleavage sites (KR/RR) (**Fig. 2.7A; Fig. S11**). These include a single copy of the polypeptide sequence QYPGGAPIGLDG (residues 58-69) and twelve copies of the polypeptide sequence QWYTG. The presence of an N-terminal glutamine (Q) residue and a C-terminal glycine (G) residue on these polypeptide sequences are indicative of post-translational modifications giving rise to a N-terminal pyro-glutaminyl (pQ) residue and a C-terminal amide group on the mature peptides (e.g. pQWYT-NH_2) based on the structure of TRH in mammals (pQHP-NH_2). Moreover, the occurrence of multiple copies of this peptide is consistent with the organisation of TRH-like precursors in vertebrates (Richter et al., 1984; Lechan et al., 1986).

TRH-type peptides are a family of neuropeptides with a widespread phylogenetic distribution indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). TRH is a hypothalamic peptide that stimulates the release of thyroid stimulating hormone (TSH) and prolactin from the anterior pituitary gland in vertebrates. In mammals, TRH stimulates release of TSH from the pituitary gland triggering release of thyroid hormones (T_3 and T_4) that stimulate metabolism in cells throughout the body and promote growth and development (Lovejoy, 2005). Moreover, TRH also acts as a neurotransmitter or neuromodulator in many other regions of the brain (Breese et al., 1981; Shibusawa et al., 2008). In amphibians and fish, however, TRH stimulates the release of pituitary growth hormone and prolactin but has little or no effect on the secretion of TSH (Galas et al., 2009). Thus, the role of TRH as a regulator of TSH release in mammals does not apply to all vertebrate species. TRH-type peptides have also been identified in a number of deuterostomian

invertebrates including the cephalochordate *B. floridae* (Jekely, 2013) and the echinoderm species *S. purpuratus* (Rowe and Elphick, 2012) and *A. japonicus* (Rowe et al., 2014). To date, TRH-like peptides have not been identified in the protostomes. However, orthologues of vertebrate TRH receptors have been identified in the ecdysozoans and lophotrochozoans (Jekely, 2013; Mirabeau and Joly, 2013). Thus, the evolutionary origin of the TRH-like neuropeptide signalling system can be traced back to the common ancestor of the bilaterians.

It is important to consider ArTRHLP in relation to TRH-like precursors/peptides identified in other phyla (**Fig. 2.7B**). The discovery of ArTRH suggests that the presence of an N-terminal pyro-glutaminyl (pQ) residue and C-terminal amide group is likely to be an ancestral feature of deuterostomian TRH-type peptides. However, it is unclear whether TRH-type peptides would have been tetrapeptides; TRH-type peptides in the vertebrates (pQHP-NH_2) and the cephalochordate *B. floridae* (pQSP-NH_2) are tripeptides (Jekely, 2013). Further insights will be provided if TRH-type peptides are identified in the hemichordates (e.g. *S. kowalevskii*), the sister phylum to the echinoderms (Rowe et al., 2014).

It is also important to consider ArTRHLP in relation to TRH-like precursors/peptides identified in other echinoderm species (**Fig. 2.7B**). ArTRHLP is the third member of the TRH-like neuropeptide family to be identified in an echinoderm after the identification of Spnp3 (SpTRHLP) in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) and Ajnp3 (AjTRHLP) in the sea cucumber *A. japonicus* (Rowe et al., 2014). Interestingly, the most abundant putative TRH-type peptide derived from SpTRHLP (pQYPG-NH_2) conforms to the pQxP motif characteristic of the human TRH peptide (pQHP-NH_2). However, the most abundant putative TRH-type peptides derived from both ArTRHLP (pQWYT-NH_2) and AjTRHLP (pQYFA-NH_2) do not share the pQxP motif found in vertebrate TRH.

However, these putative TRH-type peptides do contain an aromatic residue (tyrosine (Y); phenylalanine (F) or tryptophan (W)) in position 2 and a cyclic residue (proline (P)) in position 3, which is characteristic of vertebrate TRH-type peptides.

To date, the physiological functions of TRH-type peptides in invertebrate species and specifically in the echinoderms are currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of TRH-type peptides. In this regard, a candidate TRH-like receptor has been identified in the sea urchin *S. purpuratus* (Burke et al., 2006). However, to date, a candidate TRH-like receptor has not been identified in the starfish *A. rubens*.

(A) Thyrotropin-releasing hormone (TRH)-like precursor (ArTRHLP)

MSVTRNSGFLLVTLLFTWVVCRAELADFVENAEVAKEVSNEIEGVEAEQWQRDEDKR
QYPGGAPIGLDGKRQWYTGKRQWYTGKRDAEDSPALLAENDKRQWYTGKRSGNEEQQ
PDEANKRQWYTGKRQWYTGKRGDEDRVLDDDAVNSLKRQWYTGKRQWYTGKRSGVEQ
ADDGDLEQQYNKRQWYTGKRADDLADAADLEKRQWYTGKRQWYTGKRSDSEDEKRQW
YTGKRQWYTGRR

(B) TRH-type neuropeptide multiple sequence alignments

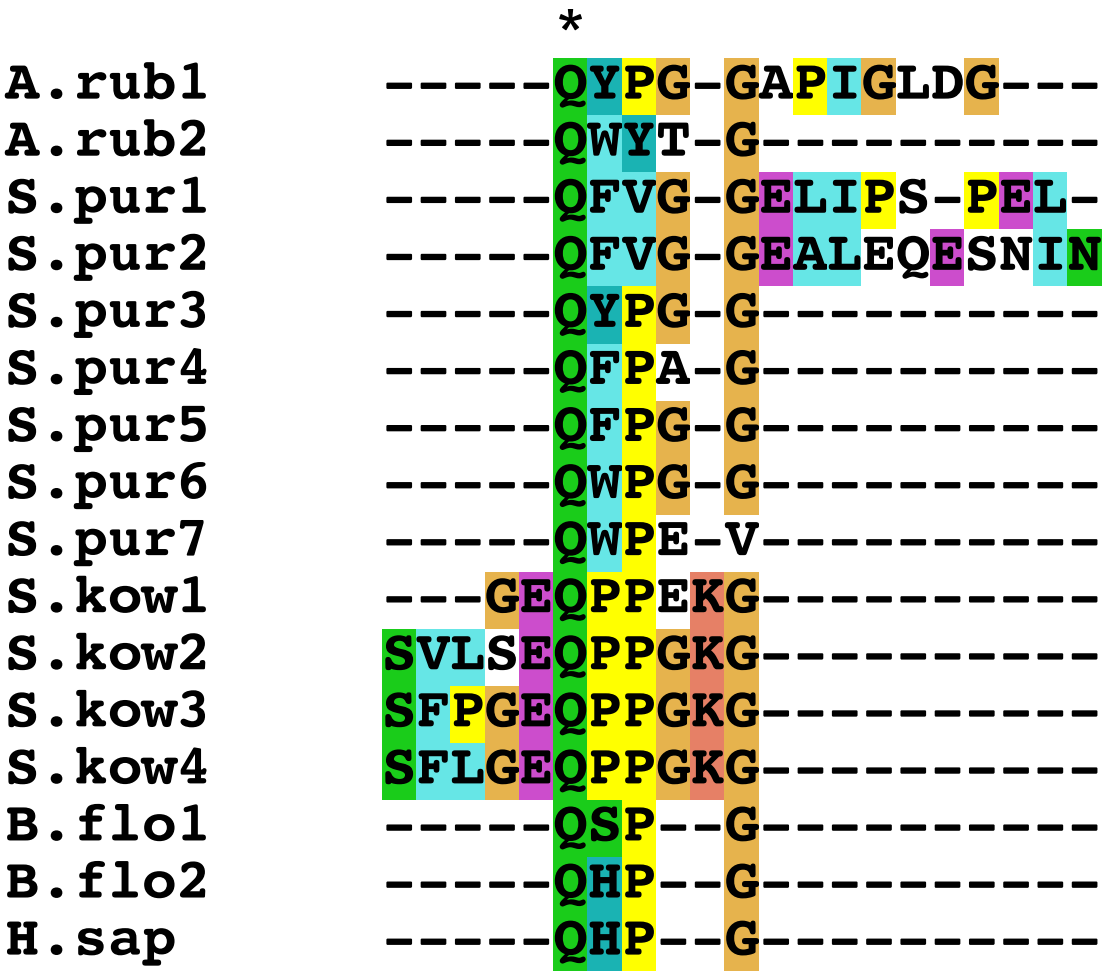


Fig. 2.7. ***A. rubens*** **TRH-like precursor (ArTRHLP). (A)** Sequence of ArTRHLPP. An N-terminal signal peptide is represented in blue, putative TRH-type neuropeptides are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative TRH-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* TRH-like precursor; *B.flo*, *Branchiostoma floridae* TRH-like precursor [GI: 260784028]; *H.sap*, *Homo sapiens* TRH-like precursor [GI: 485464565]; *S.kow*, *Saccoglossus kowalevskii* TRH-like precursor [GI: 187216047]; *S.pur*, *Strongylocentrotus purpuratus* TRH-like precursor [GI: 109402869].

2.3.2.6. TK-like precursor (ArTKLP)

The *A. rubens* tachykinin (TK)-like precursor (ArTKLP) is a 199-residue precursor protein comprising a predicted 31-residue N-terminal signal peptide and two putative TK-type peptides (**Fig. 2.8A; Fig. S12**). These include the polypeptide sequences QLWANQQSGLFG (residues 132-143) termed ArTK1 and GGGVPHVFQSGGIFG (residues 173-187) termed ArTK2 bounded by dibasic (KR) and monobasic (R) cleavage sites. The presence of an N-terminal glutamine (Q) residue on ArkTK1 and a C-terminal glycine (G) residue on both ArTK1 and ArtK2 is indicative of post-translational modifications giving rise to an N-terminal pyro-glutaminyl (pQ) residue and C-terminal amide groups on the mature peptides (e.g. pQLWANQQSGLFG-NH_2).

TK-type peptides are a family of neuropeptides with a widespread phylogenetic distribution indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). The first TK-type peptide to be discovered was substance P (Chang and Leeman, 1970; Studer et al., 1973; von Euler and Gaddum, 1931) before the subsequent discovery of neurokinin A (NKA) and neurokinin B (NKB) in mammals (Kangawa et al., 1983; Kimura et al., 1983; Nawa et al., 1984). More recently, endokinins and hemokinins have been identified as additional mammalian members of the TK-like neuropeptide family (Kurtz et al., 2002; Page et al., 2003; Zhang et al., 2000b). TK-type peptides have been implicated as neurotransmitters, neuromodulators and neurohormones in both the central and peripheral nervous system in relation to regulating intestinal motility (Shimizu et al., 2008), smooth muscle contraction (Kovac et al., 2006) and cardiovascular function (Walsh and McWilliams, 2006). TK-type peptides have been identified in a number of vertebrates (Severini et al., 2002) and have subsequently been identified in deuterostomian invertebrates including the urochordates (Satake et al., 2004) but had

not previously been identified in the ambulacrarians. TK-type peptides have also been identified in the lophotrochozoans including a number of molluscs (Kanda et al., 2003) and annelids (Ikeda et al., 1993; Kawada et al., 1999) and in the ecdysozoans including a number of arthropods (Christie et al., 1997; Nieto et al., 1998; Siviter et al., 2000; Veelaert et al., 1999) and nematodes (Nathoo et al., 2001).

It is important to consider ArTKLP in relation to TK-like precursors/peptides identified in other phyla (**Fig. 2.8B**). ArTKLP is the first member of the TK-like neuropeptide family to be identified in the ambulacrarians and specifically in an echinoderm species, with TK-type peptides so far not identified in the sea urchin *S. purpuratus* (Burke et al., 2006; Rowe and Elphick, 2012) or the sea cucumber *A. japonicus* (Rowe et al., 2014). Vertebrate TK-type peptides share the conserved C-terminal pentapeptide motif FxGLM-NH_2. Most invertebrate TK-type peptides, on the other hand, share the conserved C-terminal pentapeptide motif FxGxR-NH_2, with a few exceptions including the yellow fever mosquito *Aedes aegypti* and the octopus species *Eledone moschata* and *Octopus vulgaris*, which have vertebrate-type C-terminal pentapeptide motifs (Champagne and Ribeiro, 1994; Van Loy et al., 2010). ArTKLP contains two TK-type peptides with the C-terminal pentapeptide motifs QSGLF-NH_2 and QSGGIF-NH_2, respectively, which share the common motif GxF-NH_2, with x representing a hydrophobic residue. This motif shares more sequence similarity with the vertebrate GLM-NH_2 motif in comparison to the invertebrate GxR-NH_2 motif, with regards to the presence of hydrophobic residues. Identification of TK-type peptides in other ambulacrarian species would provide further insights into ancestral features of deuterostomian TK-type peptides.

To date, with no TK-type peptide previously identified, the physiological function of TK-type peptides in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that

mediate the effects of TK-type peptides. In this regard, candidate TK-like receptors have been identified in *S. purpuratus* (Burke et al., 2006). Importantly, two candidate TK-like receptors have also been identified in the starfish *A. rubens* and may facilitate investigation into the physiological functions of TK-type peptides in an echinoderm species (**Fig. A5J-K; Fig. S13-14**).

(A) Tachykinin (TK)-like precursor (ArTKLP)

MLLAMAPNGEMMTRFLLAAHFLLLAVSIVNARVYFNGEDETKSGLLELSEYGENEKV
DGTEDVDGQQVEDRQWKGEDQWKSGLYAAQRSLQSYPNTAKRSWPQTGMYNKQSTNW
LRALAQEPRWHSAMAKRQLWANQQSGLFGKREADMERTLPAWNVKRSAEEREFARQS
RGGGVPHVFQSGGIFGKRSSDDWAKRYE

(B) TK-type neuropeptide multiple sequence alignments

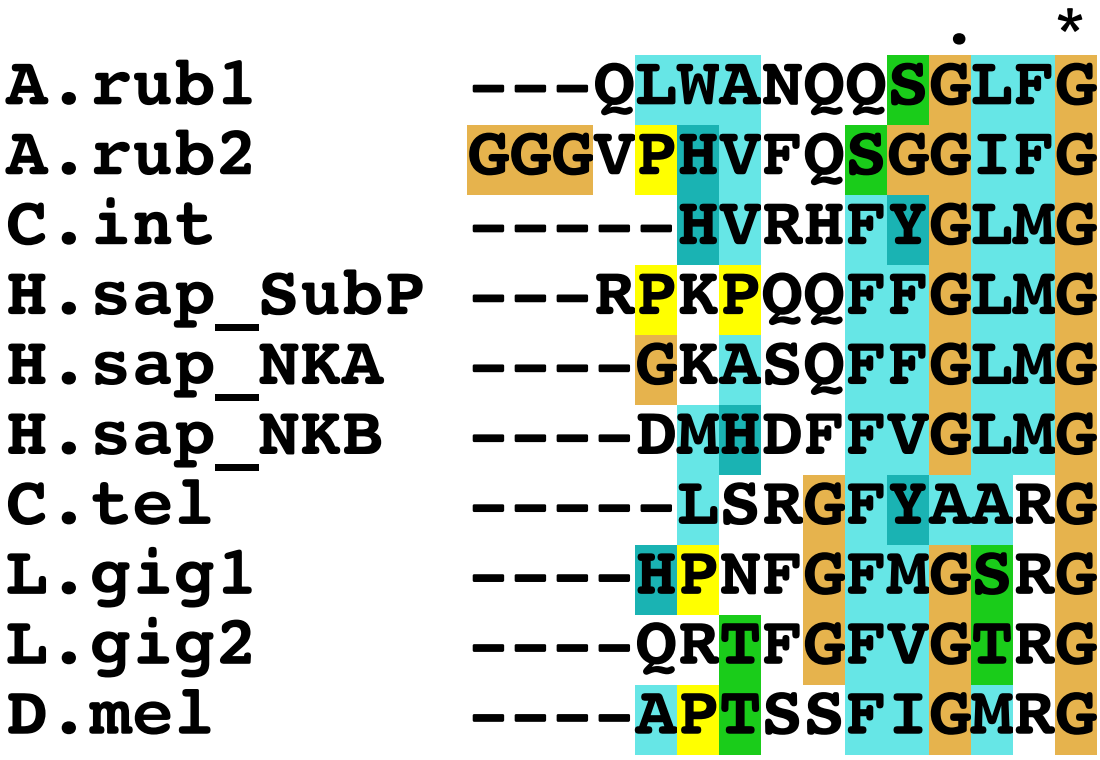


Fig. 2.8. ***A. rubens*** **TK-like precursor (ArTKLP). (A)** Sequence of ArTKLP. An N-terminal signal peptide is represented in blue, putative TK-type neuropeptides are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic/monobasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative TK-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* TK-like precursor; *C.int*, *Ciona intestinalis* TK-like precursor [GI: 74136064]; *C.tel*, *Capitella teleta* TK-like precursor [GI: 161289578]; *D.mel*, *Drosophila melanogaster* TK-like precursor [GI: 442618676]; *H.sap_SubP*, *Homo sapiens* β-prepro TK precursor [GI: 29482]; *H.sap_NKA*, *Homo sapiens* TK4 precursor [GI: 117938255]; *H.sap_NKB*, *Homo sapiens* neurokinin β precursor [GI: 48146502]; *L.gig1*, *Lottia gigantea* TK-like precursor 1 [GI: 676441944]; *L.gig2*, *Lottia gigantea* TK-like precursor 2 [GI: 163525452].

2.3.2.7. CCK-like precursor (ArCCKLP)

The *A. rubens* cholecystokinin (CCK)-like precursor (ArCCKLP) is a 163-residue precursor protein comprising a predicted 22-residue N-terminal signal peptide and two putative CCK-type peptides (**Fig. 2.9A; Fig. S15**). These include the polypeptide sequences RQSKVDDYGHGLFWG (residues 103-117) termed ArCCK1 and GGDDQYGFGLFFG (residues 139-151) termed ArCCK2 bounded by dibasic cleavage sites (KR). In ArCCK1, the presence of a tyrosine (Y) residue is indicative of a post-translational modification giving rise to a O-sulphated tyrosine (Y) residue, based the occurrence of this post-translational modification in CCK-type peptides in vertebrates (Mutt and Jorpes, 1968). The presence of a C-terminal glycine (G) residue is indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide (RQSKVDDYGHGLFW-NH_2), consistent with C-terminal amidation of CCK/gastrin-type peptides in vertebrates (Dockray et al., 1978). However, there is the possibility that the secreted bioactive peptide is formed between residues 102-117 due to the presence of a putative dibasic cleavage site (RK) at residues 100 and 101 (RRQSKVDDYGHGLFW-NH_2) or a 14-residue polypeptide formed between residues 104-117 due to the presence of a putative dibasic cleavage (RR) site at residues 100 and 101 (QSKVDDYGHGLFW-NH_2). In the latter, the presence of an N-terminal glutamine (Q) residue may be indicative of a post-translational modification giving rise to an N-terminal pyro-glutaminyl (pQ) residue on the mature peptide (pQSKVDDYGHGLFW-NH_2). Likewise, in ArCCK2, the presence of a tyrosine (Y) residue and C-terminal phenylalanine (F) and glycine (G) residues are indicative of post-translational modifications giving rise to an O-sulphated tyrosine (Y) residue and a C-terminal amide group on the mature peptide respectively (GGDDQYGFGLFF-NH_2).

CCK-type peptides are a family of neuropeptides with a widespread

phylogenetic distribution indicative of an ancestral bilaterian origin (Hewes and Taghert, 2001; Janssen et al., 2008; Jekely, 2013; Mirabeau and Joly, 2013). CCK is structurally and functionally related to gastrin; both peptides are characterised by the C-terminal tetrapeptide WMDF-NH_2, which has been shown to be necessary for biological activity and is suggestive of a common evolutionary origin (Dockray, 1977; Johnsen, 1998; Larsson and Rehfeld, 1977). CCK has numerous regulatory functions in both the gastrointestinal system and in the brain of mammals. In the gastrointestinal system, CCK/gastrin have been implicated in functions including gut development, gallbladder contraction, gastrointestinal motility and pancreatic secretion of digestive enzymes (Guilloteau et al., 2006). In the brain, CCK/gastrin have been implicated in learning and memory, angiogenesis, nociception and regulating food intake (Dufresne et al., 2006; Johnsen, 1998; Konturek et al., 2003). CCK-type peptides have subsequently been identified in other vertebrates (Johnsen, 1998) and deuterostomian invertebrates including the urochordate *C. intestinalis* (Johnsen and Rehfeld, 1990), the hemichordate *S. kowalevskii* (Mirabeau and Joly, 2013) and the echinoderm *S. purpuratus* (Mirabeau and Joly, 2013). Interestingly, however, CCK-type peptides and receptors appear to have been lost in the cephalochordate *B. floridae* (Mirabeau and Joly, 2013).

CCK-type peptides have been found to be orthologous to sulfakinin (SK)-type peptides (Schoofs and Nachman, 2006; Schoofs et al., 2013). SK was first isolated from head extracts of the Madeira cockroach *Leucophaea maderae* and displays myotropic activity on the hindgut (Nachman et al., 1986a; Nachman et al., 1986b). SK-type peptides have been implicated in numerous regulatory functions in multiple insect species including myotropic effects on the hindgut and foregut (Maestro et al., 2001), heart (Nichols et al., 2009) and body wall muscles (Chen et al., 2012). SK-type peptides have also been shown to involved in regulating food

intake in multiple insect species, including the desert locust *Schistocerca gregaria* (Wei et al., 2000) and regulating digestive enzyme release in the beetle *Rhynchophorus ferrugineus* (Nachman et al., 1997) and the moth *Opisina arenosella* (Harshini et al., 2002). SK-type peptides are characterised by the C-terminal heptapeptide DYGHMRF-NH_2, which is often sulphated at the tyrosine (Y) residue (Schoofs et al., 2013). SK-type peptides have been identified in the ecdysozoans including in a number of arthropods (Schoofs et al., 2013) and nematodes (Janssen et al., 2008) and in the lophotrochozoans (Mirabeau and Joly, 2013). It would therefore appear that the CCK/gastrin/SK-like neuropeptide system can be traced back to the common ancestor of the bilaterians and is likely to have an ancient role in regulating physiological processes in the gastrointestinal system.

It is important to consider ArCCKLP in relation to CCK/gastrin/SK-like precursors/peptides identified in other phyla (**Fig. 2.9B**). Most vertebrate CCK-type peptides typically share the conserved C-terminal octapeptide motif DYMGWMDF-NH_2. Most SK-type peptides, on the other hand, typically share the conserved C-terminal heptapeptide motif DYGHMRF-NH_2. ArCCKLP contains two CCK-type peptides with the C-terminal octapeptide motifs DYGHGLFW-NH_2 and (R)QYGFGLFF-NH_2, respectively, which share the common motif x_1YGx_2GLFx_3-NH_2, with x_3 representing a shared hydrophobic residue. This motif shares sequence similarity with both the vertebrate and SK-type motifs, including the likely presence of a conserved O-sulphated tyrosine (Y) residue and an amidated phenylalanine (F) residue (with the exception of ArCCK1, which has a tryptophan (W) residue).

It is also important to consider ArCCKLP in relation to CCK/gastrin-like precursors/peptides identified in other echinoderm species (**Fig. 2.9B**). ArCCKLP is only the second member of the CCK-like neuropeptide family to be identified in an echinoderm after the identification of a CCK-like precursor (SpCCKLP) in the sea

urchin *S. purpuratus* (Mirabeau and Joly, 2013). To date, a CCK-like precursor has not been identified in a transcriptome dataset from the sea cucumber *A. japonicus* (Rowe et al., 2014). SpCCKLP has two CCK-type peptides with the C-terminal octapeptides DYGHGMFF-NH_2 and DYNWGMWF-NH_2, respectively, which share the common motif $DYx_1x_2Gx_3F$-NH_2, with x_3 representing a shared hydrophobic residue. This motif shares similarity with the ArCCKLP motif, including the presence of a conserved glycine (G) residue in comparison to the vertebrate and SK-type motifs respectively.

To date, the physiological function of CCK-type peptides in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of CCK-type peptides. In this regard, a candidate CCK-like receptor has been identified in the sea urchin *S. purpuratus* (Burke et al., 2006). Importantly, a candidate CCK-like receptor has also been identified in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of CCK-type peptides in an echinoderm species (**Fig. A5A; Fig. S16**).

(A) Cholecystokinin (CCK)-like precursor (ArCCKLP)

MSSWLTVAIATVTCLLLSPITCLPLHDVADGKERRELLHSTWLDPSGSTGQGTEELA
ETSKRLLGDNNRDSGIIDLLVALRDTNTNPRDLYLHGNTETARKRRQSKVDDYGHGL
FWGKRGSNWSDHGVRAMTDKDTKRGGDDQYGFGLFFGKRNEEDYEDFTL

(B) CCK-like neuropeptide multiple sequence alignments

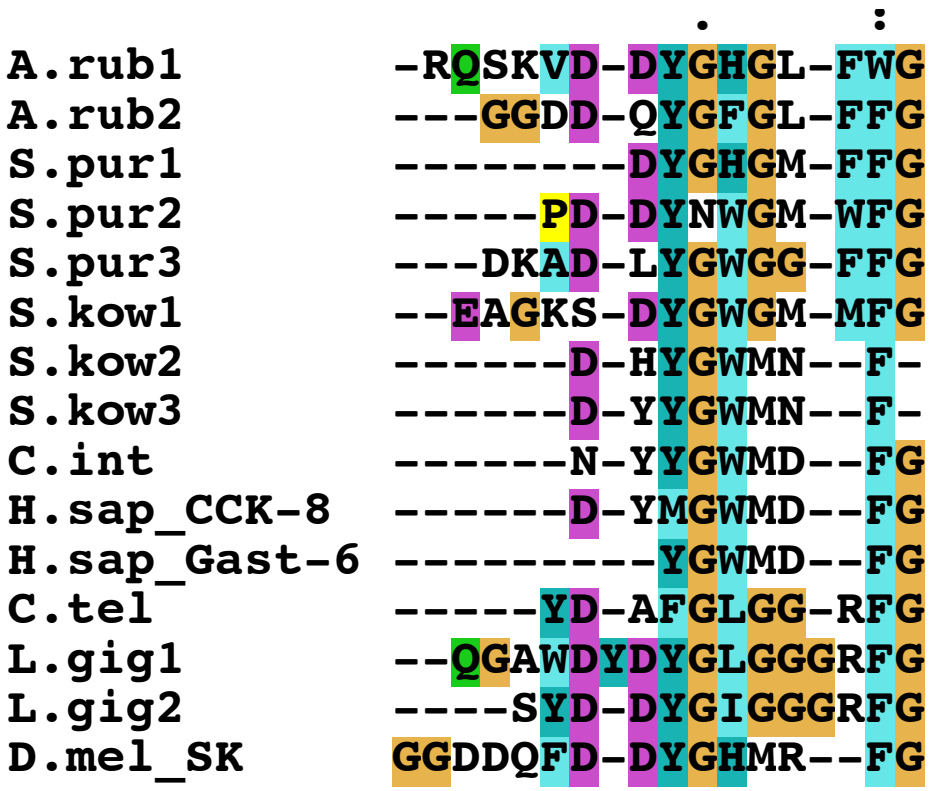


Fig. 2.9. *A. rubens* CCK-like precursor (ArCCKLP). (A) Sequence of ArCCKLP. An N-terminal signal peptide is represented in blue, putative CCK-type neuropeptides are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative CCK-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* CCK-like precursor; *C.int*, *Ciona intestinalis* CCK-like precursor [GI: 10799472]; *C.tel*, *Capitella teleta* CCK-like precursor [GI: 161296032]; *D.mel_SK*, *Drosophila melanogaster* sulfakinin (SK) precursor [GI: 386765036]; *H.sap_CCK-8*, *Homo sapiens* CCK-8 precursor [GI: 30582820]; *H.sap_Gast-6*, *Homo sapiens* gastrin precursor [GI: 47481291]; *L.gig1*, *Lottia gigantea* CCK-like precursor 1 [GI: 161296032]; *L.gig2*, *Lottia gigantea* CCK-like precursor 2 [GI: 52414496]; *S.kow1*, *Saccoglossus kowalevskii* CCK-like precursor 1 [GI: 585688033]; *S.kow2*, *Saccoglossus kowalevskii* CCK-like precursor 2 [GI: 187061456]; *S.pur*, *Strongylocentrotus purpuratus* CCK-like precursor [GI: 390355380].

2.3.2.8. OX-like precursors (ArOXLP1-2)

The *A. rubens* orexin (OX)-like precursor 1 (ArOXLP1) is a 112-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide followed by a putative OX-type peptide with the predicted polypeptide sequence SNADSACCARTFRCNLRSDCTCMVREILCRDPSEGMLNSG (residues 25-64) termed ArOX1 bounded by a dibasic cleavage site (KR) (**Fig. 2.10A; Fig. S17**). The presence of a C-terminal glycine (G) residue is indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide (SNADSACCARTFRCNLRSDCTCMVREILCRDPSEGMLNS-NH_2). It is noteworthy that this putative OX-type peptide contains six cysteine (C) residues (residues 31, 32, 38, 44, 46 and 53), which are likely to form up to three disulphide bridges in accordance with OX-type peptides in vertebrates that form two disulphide bridges (Sakurai et al., 1998).

The *A. rubens* OX-like precursor 2 (ArOXLP2) is a 161-residue precursor protein comprising a predicted 26-residue N-terminal signal peptide followed by a putative OX-type peptide with the predicted polypeptide sequence NACCRGTCHDIPPGCNCPYKSYLCGELNALTMG (residues 27-59) termed ArOX2 bounded by a dibasic cleavage site (KR) (**Fig. 2.10B; Fig. S18**). The presence of a C-terminal glycine (G) residue is indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide (NACCRGTCHDIPPGCNCPYKSYLCGELNALTMG-NH_2). It is noteworthy that this putative OX-type peptide contains six cysteine (C) residues (residues 29, 30, 34, 41, 43 and 50), which are likely to form up to three disulphide bridges in accordance with OX-type peptides in vertebrates that form two disulphide bridges (Sakurai et al., 1998).

OX-type peptides are a family of neuropeptides with a widespread phylogenetic distribution indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). The OX's are hypothalamic neuropeptides that were first discovered as regulators of feeding behaviour in mammals (de Lecea et al., 1998; Sakurai et al., 1998). However, OX's have subsequently been implicated as key regulators of the processes of sleep and wakefulness in mammals (Sakurai, 2007). Similarly, OX's have also been shown to regulate feeding behaviour and the processes of sleep and wakefulness in teleost fish (Matsuda et al., 2012). OX-type peptides have subsequently been identified in a number of deuterostomian invertebrates including the cephalochordate *B. floridae* (Jekely, 2013; Mirabeau and Joly, 2013), the hemichordate *S. kowalevskii* (Jekely, 2013; Mirabeau and Joly, 2013) and the echinoderm *S. purpuratus* (Jekely, 2013; Mirabeau and Joly, 2013).

OX-type peptides have recently been found to be orthologous to allatotropin (AT)-type peptides based on receptor orthology (Jekely, 2013; Mirabeau and Joly, 2013). AT was first isolated through its regulatory function in stimulating the synthesis and secretion of juvenile hormone (JH) from the *corpora allata* and has subsequently been implicated in regulating cardioacceleration and the inhibition of active ion transport in the midgut of the tobacco hornworm *M. sexta* (Kataoka et al., 1989; Veenstra et al., 1994). AT-type peptides have therefore been identified in the ecdysozoans (Elekonich and Horodyski, 2003; Weaver and Audsley, 2009) and in the lophotrochozoans including molluscs (Harada et al., 1993; Li et al., 1993; Veenstra, 2010) and annelids (Ukena et al., 1995; Veenstra, 2011). Interestingly, however, OX-type peptides and receptors appear to have been lost in the nematode *C. elegans* (Jekely, 2013; Mirabeau and Joly, 2013).

It is important to consider ArOXLP1 and ArOXLP2 in relation to OX-like precursors/peptides identified in other phyla (**Fig. 2.10C**). OX-type peptides are

highly conserved in the vertebrates with two OX-type peptides (termed OX-A and OX-B) encoded in tandem on one OX-like precursor protein (Tsujino and Sakurai, 2009). However, two separate OX-like precursors have been identified in the starfish *A. rubens*, which is consistent with other OX-like precursors in the ambulacrarians (Jekely, 2013; Mirabeau and Joly, 2013). OX-A in vertebrates has an N-terminal pyro-glutaminyl (pQ) residue and four cysteine (C) residues, which has been shown to form two disulphide bridges (Sakurai et al., 1998). ArOX1 and ArOX2 have six cysteine (C) residues suggesting the presence of three disulphide bridges; however, both ArOX1 and ArOX2 do not contain an N-terminal pyro-glutaminyl (pQ) residue. Interestingly, however, only one OX-like precursor has been identified in the hemichordate *S. kowalevskii* (Jekely, 2013; Mirabeau and Joly, 2013). The *S. kowalevskii* OX-like precursor contains one OX-type peptide with an N-terminal pyro-glutaminyl (pQ) residue and has six cysteine (C) residues suggesting the presence of three disulphide bridges. Thus, the *S. kowalevskii* OX-type peptide shares features of both OX-A and ArOX1 and ArOX2, consistent with the phylogenetic position of the hemichordates as a sister group to the echinoderms. Furthermore, these findings indicate that the occurrence of six cysteine (C) residues may be a characteristic of OX-type peptides in the ambulacrarians.

It is also important to consider ArOXLP1 and ArOXLP2 in relation to OX-like precursors/peptides identified in other echinoderm species (**Fig. 2.10C**). ArOXLP1 and ArOXLP2 are only the second members of the OX-like neuropeptide family to be identified in an echinoderm after the identification of two OX-like precursors (SpOXLP1-2) in the sea urchin *S. purpuratus* (Jekely, 2013; Mirabeau and Joly, 2013). Similarly to the starfish *A. rubens*, there are two OX-like precursors each containing one OX-type peptide in the sea urchin *S. purpuratus*; both OX-type peptides have six cysteine (C) residues suggesting the presence of three disulphide

bridges. Interestingly, however, the OX-type peptide derived from SpOXLP1 has a putative N-terminal pyro-glutaminyl (pQ) residue, which is consistent with vertebrate OX-A and an OX-type peptide in the hemichordate *S. kowalevskii*.

To date, the physiological function of OX-type peptides in invertebrate species is not particularly well understood and in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of OX-type peptides. In this regard, two candidate OX-like receptors have been identified in the sea urchin *S. purpuratus* (Burke et al., 2006). Importantly, three candidate OX-like receptors have also been identified in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of OX-type peptides in an echinoderm species (**Fig. A5G-I; Fig. S19-21**).

(A) Orexin (OX)-like precursor 1 (ArOXLP1)

MMTRSYLVLCCLVIWAFLATGTLPSNADSACCARTFRCNLRSDCTCMVREILCRDPS
EGMLNSGKRSPPSDRQQQQLNVLESSGYHANPRMMPSSSKYRRALPLATMNEDFL

(B) OX-like precursor 2 (ArOXLP2)

MRTTTLFVIQTVVVIGYFTCSSTAAANACCRGTCHDIPPGCNCPYKSYLCGELNALT
MGKRKADDTSYLLTQEQETQQQNQQRRTQQTQPWVDRQPDDDRIVDVLNNLLKLFKE
THQGDQDGFDLQDQSDDWEPVTSSRKQQSENAHNVYRHQPLFSADIL

(C) OX-like neuropeptide multiple sequence alignments

```
                                                          .     *
A.rub1      SNADSACCARTFRCNLRSDCTCMVREILCRD--PSEGMLNSG
A.rub2      ----NACCR-GTCHDIPPGCNCPYKSYLCG----ELNALTMG
S.pur1      ---QSPCCRRAKGCSFPPGCHCPLKMSFCGD--PSRGLQIVG
S.pur2      ---DRACCKRTVGCNLRSDCTCRIREITCTD--PSLGLQNYG
S.kow       ---QPQCCR-GVGCKIPPNCKCPFQSIICDN--PTKNVLTAG
B.flo1      ---FPRCCW-------RRTCNCPHFRRLFGPRNHGHGILTVG
B.flo2      --QKPACCE-------SGSCICPYVSRLHSPGNHGYGILTTG
H.sap_OXA   -QPLPDCCR-------QKTCSCRLYELLHGAGNHAAGILTLG
H.sap_OXB   -RSGPPGLQ-------G-----RLQRLLQASGNHAAGILTMG
```

Fig. 2.10. *A. rubens* OX-like precursors. (A) Sequence of ArOXLP1. **(B)** Sequence of ArOXLP2. N-terminal signal peptides are represented in blue, putative OX-type neuropeptides (with cysteine (C) residues underlined) are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green. **(C)** ClustalX v.2.1 multiple sequence alignment of putative OX-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub1*, *Asterias rubens* OX-like precursor 1; *A.rub2*, *Asterias rubens* OX-like precursor 2; *B.flo1*, *Branchiostoma floridae* OX-like precursor 1 [GI: 260807454]; *B.flo2*, *Branchiostoma floridae* OX-like precursor 2 [GI: 260780718]; *H.sap_OXA-B*, *Homo sapiens* OX (hypocretin) precursor [GI: 4557634]; *S.kow*, *Saccoglossus kowalevskii* OX-like precursor [GI: 585662697]; *S.pur1*, *Strongylocentrotus purpuratus* OX-like precursor 1 [GI: 346420309]; *S.pur2*, *Strongylocentrotus purpuratus* OX-like precursor 2 [GI: 346419879].

2.3.2.9. LQ-like precursor (ArLQLP)

The *A. rubens* luqin (LQ)-like precursor (ArLQLP) is a 106-residue precursor protein comprising a predicted 44-residue N-terminal signal peptide followed by a putative LQ-type peptide with the predicted peptide sequence EEKTRFPKFMRWG (residues 45-57) termed ArLQ bounded by a dibasic cleavage site (KR) (**Fig. 2.11A; Fig. S22**). The presence of a C-terminal glycine (G) residue is indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide (EEKTRFPKFMRW-NH_2).

LQ-type peptides are a family of neuropeptides with a phylogenetic distribution recently attributed to be indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). LQ was first identified in the California sea slug *Aplysia californica* and named on account of the expression of its precursor protein in the dorsal left upper quadrant (LUQ) cells of the abdominal ganglion (Aloyz and DesGroseillers, 1995). Prior to the discovery of LQ in *A. californica*, a closely related neuropeptide termed *Achatina* cardioexcitatory peptide-1 (or ACEP-1) was isolated from the African giant snail *Achatina fulica* on account of its excitatory effect on the heart and other muscles (Fujimoto et al., 1990). Subsequently, a closely related peptide termed *Lymnaea* cardioexcitatory peptide (or LyCEP) that also has a cardioexcitatory effect was isolated from the pond snail *Lymnaea stagnalis* (Tensen et al., 1998). LQ-type peptides have also been identified in a number of lophotrochozoans including other molluscs (Veenstra, 2010) and annelids (Veenstra, 2011) and have recently been identified as orthologous to the RYamides identified in the ecdysozoans, including a number of arthropods and nematodes (Jekely, 2013; Mirabeau and Joly, 2013). LQ-type peptides have also been identified in a number of deuterostomian invertebrates including the hemichordate *S. kowalevskii* and the echinoderm species *S. purpuratus* (Jekely,

2013; Mirabeau and Joly, 2013) and *A. japonicus* (Rowe et al., 2014). Interestingly, however, LQ-type peptides or receptors have not yet been identified in the chordates, suggesting that the LQ-like neuropeptide signalling system may have been lost in the chordate lineage.

It is important to consider ArLQLP in relation to LQ-like precursors/peptides identified in other phyla (**Fig. 2.11B**). ArLQLP (EEKTRFPKFMRW-NH_2) contains one LQ-type peptide, similar to the LQ-type peptides in the lophotrochozoans and *S. kowalevskii*, as opposed to the RYamide-type peptides in the ecdysozoans that are repetitive, which is likely a derived feature (Jekely, 2013). ArLQ has a C-terminal RW-NH_2 motif, which shares the common second hydrophobic residue with both the RF-NH_2 motif in the lophotrochozoans and the RY-NH_2 motif in the ecdysozoans. Moreover, the presence of two cysteine (C) residues typically separated by a 10-residue polypeptide sequence in the C-terminal region of the ArLQLP precursor appears to be a conserved feature of all LQ-like precursors (**Fig. 2.11B**) (Jekely, 2013; Mirabeau and Joly, 2013).

It is also important to consider ArLQLP in relation to LQ-like precursors/peptides identified in other echinoderm species (**Fig. 2.11B**). Comparison of ArLQ (EEKTRFPKFMRW-NH_2) with LQ-type peptides in the sea urchin *S. purpuratus* (PHKFMRW-NH_2) and the sea cucumber *A. japonicus* (PYKFMRW-NH_2) reveals high sequence similarity, sharing the C-terminal KFMRW-NH_2 motif. Moreover, similarly to LQ-type peptides in the echinoderm species *S. purpuratus* and *A. japonicus*, ArLQLP has two cysteine (C) residues separated by a 10-residue polypeptide sequence in the C-terminal region of the precursor.

To date, the physiological function of LQ-type peptides in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of LQ-type peptides. In this

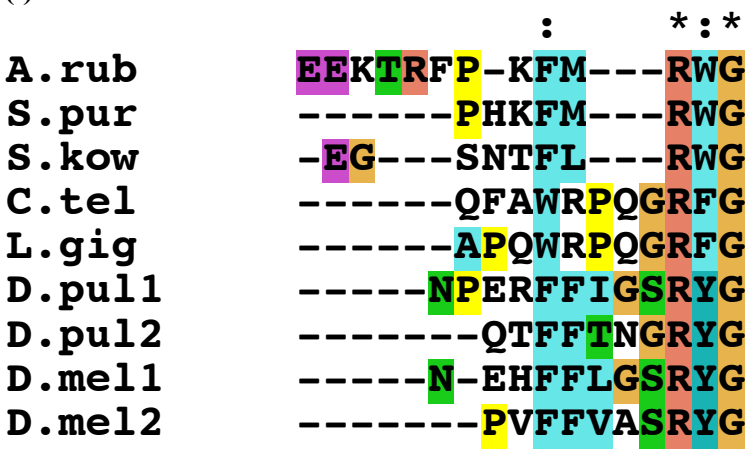
regard, two candidate LQ-like receptors orthologous to the LyCEP receptor (Tensen et al., 1998) have been identified in the sea urchin *S. purpuratus* (Jekely, 2013; Mirabeau and Joly, 2013). Similarly, an ortholog of the LyCEP receptor has also been identified in the sea cucumber *A. japonicus* (Rowe et al., 2014). Importantly, two candidate LyCEP-type receptors have been identified here in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of LQ-type peptides in an echinoderm species (**Fig. A5D-E; Fig. S23-24**).

(A) Luqin (LQ)-like precursor (ArLQP)

MTNRTAQEQCTRAGSICPSIIRFSAWLLLTILVAQVLLGTTAKAEEKTRFPKFMRWG
KRYSPDYVVMDDNELKDEMKLPVFGNGEVLCKNVASGGLYRCGKVPATA

(B) LQ-type neuropeptide/precursor multiple sequence alignments

(i)



(ii)

Fig. 2.11. *A. rubens* LQ-like precursor (ArLQP). (A) Sequence of ArLQP. An N-terminal signal peptide is represented in blue, a putative LQ-type neuropeptide is represented in red, a C-terminal glycine (G) residue is represented in green and dibasic/monobasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of **(i)** putative LQ-type neuropeptides and **(ii)** a C-terminal domain of LQ-like precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* LQ-like precursor; *C.tel*, *Capitella teleta* LQ-like precursor [GI: 161280144]; *D.mel*, *Drosophila melanogaster* RY-amide precursor [GI: 442634438]; *D.pul*, *Daphnia pulex* LQ-like precursor [JGI: 251691]; *L.gig*, *Lottia gigantea* LQ-like precursor [GI: 163510328]; *S.kow*, *Saccoglossus kowalevskii* LQ-like precursor [GI: 585716464]; *S.pur*, *Strongylocentrotus purpuratus* LQ-like precursor [GI: 390331828].

2.3.2.10. KP-like precursor (ArKPLP)

The *A. rubens* kisspeptin (KP)-like precursor (ArKPLP) is a 149-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide and two putative KP-type peptides (**Fig. 2.12A; Fig. S25**). These include the peptide sequences SGRCRSGTKCIMRGPNPNTASRVLPFG (residues 87-113) termed ArKP1 and GRGPPKNSRARGGRTLLPFG (residues 127-146) termed ArKP2 bounded by dibasic cleavage sites (KR). The presence of C-terminal glycine (G) residues on each of these sequences are indicative of post-translational modifications giving rise to a C-terminal amide group on the mature peptides (e.g. GRGPPKNSRARGGRTLLPF-NH_2). It is noteworthy that ArKP1 contains two cysteine (C) residues (residues 90 and 96), which suggest the presence of a disulphide bridge in the mature peptide.

KP-type peptides are a family of neuropeptides with a phylogenetic distribution recently attributed to be indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). KP (or KiSS1) is a hypothalamic peptide that was originally discovered in humans as a metastasis-suppressor gene (Harms et al., 2003; Lee et al., 1996). However, subsequently, KP has been shown to have a key role in reproductive functions in the hypothalamic-pituitary axis in mammals (Oakley et al., 2009). KP has been shown to regulate the activity of GnRH neurons both directly (Colledge, 2009) and indirectly (Pielecka-Fortuna et al., 2008; Zhang et al., 2008) whilst KP has also been shown to act directly on gonadotropes (Richard et al., 2009). Similarly, non-mammalian vertebrate KP-type peptides have also been implicated in the regulation of reproductive function in various fish (Elizur, 2009; Oakley et al., 2009; Zohar et al., 2010). However, KP expression outside of the hypothalamic-pituitary axis has also been shown in mammals. For example, KP is expressed in the hippocampus and has been implicated with various neurological processes including cognition and neurogenesis (Arai, 2009). KP-type peptides have

subsequently been identified in a number of deuterostomian invertebrates including the cephalochordate *B. floridae* (Mirabeau and Joly, 2013) and, recently, in the echinoderm *S. purpuratus* (Elphick and Mirabeau, unpublished data). Interestingly, there has been a large expansion of KP-like receptors in the cephalochordate *B. floridae* and in the ambulacrarians (Mirabeau and Joly, 2013). KP-like receptors have also been identified in the lophotrochozoa but, to date, a KP-type peptide has yet to be identified. However, KP-type peptides and receptors appear to have been lost in lineages leading to the urochordates and the ecdysozoans (Mirabeau and Joly, 2013).

It is important to consider ArKPLP in relation to KP-like precursors/peptides identified in other phyla (**Fig. 2.12B**). ArKPLP contains two KP-type peptides – ArKP1 (SGRCRSGTKCIMRGPNPNTASRVLPF-NH_2) and ArKP2 (GRGPPKNSRARGGRTLLPF-NH_2) – whilst vertebrate KP-like precursors contain one KP-type peptide. ArKP1 and ArKP2 share the common C-terminal motif LxF-NH_2, which is a characteristic feature of KP-type peptides. However, ArKP1 has two cysteine (C) residues, suggesting the presence of a disulphide bridge or a homodimeric protein formed by up to two intermolecular disulphide bridges, which is not a feature of KP-type peptides in other phyla.

It is also important to consider ArKPLP in relation to KP-like precursors/peptides identified in other echinoderms (**Fig. 2.12B**). ArKPLP is only the second KP-like precursor to be identified in the ambulacrarians and specifically in an echinoderm after the identification of a KP-like precursor (SpKPLP) in the sea urchin *S. purpuratus* (Elphick and Mirabeau, unpublished data). SpKPLP has two KP-type peptides – SpKP1 (SRCRGRQCRNVGGLNPNANLRPLPF-NH_2**)** and SpKP2 (GRTKNRIRERVPHFLPF-NH_2**)**. Both SpKP1 and SpKP2 (like ArKP1 and ArKP2) share the common C-terminal motif LxF-NH_2 characteristic of KP-type

peptides. Moreover, and similarly to ArKP1, SpKP1 also has two cysteine (C) residues suggesting the presence of a disulphide bridge, or a homodimeric protein formed by up to two intermolecular disulphide bridges – this may be a characteristic feature of KP-type peptides in the echinoderms.

To date, the physiological function of KP-type peptides in invertebrate species is unknown. Therefore, the discovery of ArKPLP (and SpKPLP) are important findings. Investigation of the physiological roles of KP-type peptides in the echinoderms would be facilitated by the identification of receptors that mediate their effects. In this regard, four candidate KP-like receptors have been identified in the sea urchin *S. purpuratus* (Mirabeau and Joly, 2013). Importantly, nine candidate KP-like receptors have also been identified in the starfish *A. rubens*, which confirms a large expansion of KP-like receptors in the ambulacrarians and may facilitate investigation into the physiological functions of KP-type peptides in an echinoderm species (**Fig. A6A-I; Fig. S26-34**).

(A) Kisspeptin (KP)-like precursor (ArKPLP)

MEWFTKCCLLVILAVCFGSSFVLGDGRNLQGYNGDLYNGEFENEEGTEALRNIIGQI
IDDVDAKNNIRTAILEDTLEHAQYEPDKRSGRCRSGTKCIMRGPNPNTASRVLPFGK
REDDSPNKLARRGRGPPKNSRARGGRTLLPFGKRR

(B) KP-type neuropeptide multiple sequence alignments

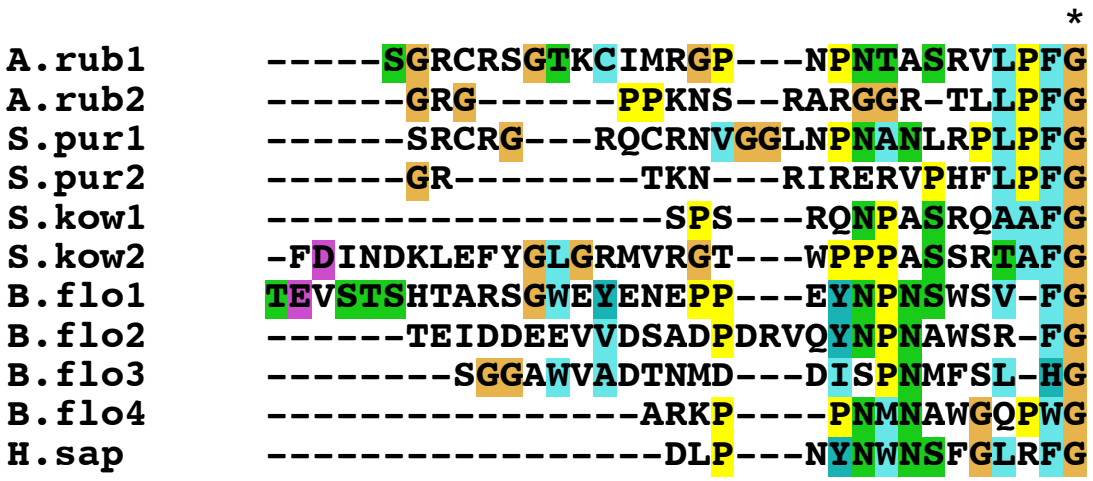


Fig. 2.12. ***A. rubens*** **KP-like precursor (ArKPLP). (A)** Sequence of ArKPLP. An N-terminal signal peptide is represented in blue, putative KP-type neuropeptides (with cysteine (C) residues underlined) are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative KP-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* KP-like precursor; *B.flo1*, *Branchiostoma floridae* KP-like precursor 1 [GI: 260826607]; *B.flo2*, *Branchiostoma floridae* KP-like precursor 2 [GI: 260827077]; *B.flo3*, *Branchiostoma floridae* KP-like precursor 3 [GI: 260826605]; *B.flo4*, *Branchiostoma floridae* KP-like precursor 4 [GI: 260793233]; *H.sap*, *Homo sapiens* KiSS-1 metastasis-suppressor precursor [GI: 21950713]; *S.kow*, *Saccoglossus kowalevskii* KP-like precursor [GI: 187123982]; *S.pur*, *Strongylocentrotus purpuratus* KP-like precursor [GI: 374768013].

2.3.2.11. SMS-like precursor (ArSMSLP)

The *A. rubens* somatostatin (SMS)-like precursor (ArSMSLP) is a 132-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide and a putative SMS-type peptide with the predicted polypeptide sequence KCIGRFQPFSMPC (residues 120-132) termed ArSMS bounded by a dibasic cleavage site (KR) (**Fig. 2.13A; Fig. S35**). It is noteworthy that this putative SMS-type peptide contains two cysteine (C) residues, which may form an intramolecular disulphide bridge in accordance with SMS-type peptides (Rivier et al., 1975). ArSMSLP was identified through sequence similarity with Spnp19 in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012). However, upon the discovery of ArSMSLP and alignments with SMS-like precursors (**Fig. 2.13B**), it is likely that Spnp19 is a SMS-like precursor comprising a putative SMS-type peptide with the predicted polypeptide sequence GKCMGRFGPYMLNC, bounded by a monobasic (R) and dibasic cleavage site (KR).

SMS-type peptides are a family of neuropeptides with a phylogenetic distribution recently attributed to be indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013; Veenstra, 2009). SMS was first characterised from the sheep hypothalamus (Brazeau et al., 1973) and was initially described as a neuroendocrine peptide regulating pituitary hormone release (Epelbaum, 1986; Fodor et al., 2005). An SMS-related peptide was subsequently discovered in *H. sapiens* termed cortistatin (de Lecea et al., 1997) and has been shown to be present throughout tetrapod vertebrates (Tostivint et al., 2008). Interestingly, however, additional SMS-type precursors termed SMS 3–6 are also present in teleost fish (Liu et al., 2010). SMS-type peptides have subsequently been implicated in a number of physiological functions including neuromodulation, motor activity, sensory processing and cognition (Viollet et al., 2008). SMS-type peptides

have been identified in other vertebrates (Ocampo Daza et al., 2012) and a candidate SMS-type peptide has recently been identified in the cephalochordate *B. floridae*, revealing that SMS-type peptides occur throughout the chordates (Mirabeau and Joly, 2013).

SMS-type peptides have recently been found to be orthologous to allatostatin-C (AST-C) (Jekely, 2013; Mirabeau and Joly, 2013; Veenstra, 2009). AST's are a family of peptides that inhibit JH biosynthesis in the *corpora allata*. AST's were first isolated from the cockroach *Diploptera punctata* and these were termed AST-A (Pratt et al., 1991; Pratt et al., 1989; Woodhead et al., 1989). Neuropeptides with allastatic activity identified from the cricket *Gryllus bimaculatus* were subsequently termed AST-B (Lorenz et al., 1995) and from the tobacco hornworm *M. sexta* were subsequently termed AST-C despite structural differences to AST-A-type peptides (Kramer et al., 1991). AST-C-type peptides have been identified in a number of arthropods including numerous insect species (Li et al., 2006; Li et al., 2008b; Price et al., 2002; Williamson et al., 2001) and also in the tick *Ixodes scapularis* (Veenstra, 2009). Interestingly, it has recently been reported that these species also have precursors comprising a second AST-C-type peptide termed AST-CC (Veenstra, 2009).

It is important to consider ArSMSLP in relation to SMS-like precursors/peptides that have been identified in other phyla (**Fig. 2.13B**). It is likely that Spnp19 identified in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) is in fact an SMS-like precursor based on the discovery of ArSMS and further support from neuropeptide alignments. ArSMSLP comprises a putative 13-residue SMS-type peptide (KCIGRFQPFSMPC) sharing high sequence similarity with the putative 14-residue SMS-type peptide (GKCMGRFGPYMLNC) derived from Spnp19 (Rowe and Elphick, 2012). Both ArSMS and the SMS-type peptide derived from Spnp19

have a GRF motif and two cysteine (C) residues, which is a characteristic of SMS/cortistatin-type peptides in the vertebrates. However, the SMS precursor in *H. sapiens* contains two biologically active SMS peptides derived from differential cleavage of the SMS precursor – the 14-residue SMS-14 and the 28-residue SMS-28 (Pradayrol et al., 1980).

To date, the physiological roles of SMS-type peptides in deuterostomian invertebrates and more specifically in the echinoderms are currently unknown. To this extent, investigation would be facilitated by identification of receptors that mediate the effects of SMS-type peptides. Importantly, three candidate SMS-like receptors have been identified in the starfish *A. rubens* and may facilitate investigation into the physiological functions of SMS-type peptides in an echinoderm species (**Fig. A6L-N; Fig. S36-38**).

(A) Somatostatin (SMS)-like precursor (ArSMSLP)

MRFCCVVVVLPLICVLAGCLVAHAAPRRGGGGNSDPRWKRNFSPPGMQSSGGSYNKG
DLVERILNRLQERLLGKVDLSQTNTWHGNQSPKELDLQRYSDQEDEFIDDDDEVPNR
PAIKRKCIGRFQPFSMPC

(B) SMS-type neuropeptide multiple sequence alignments

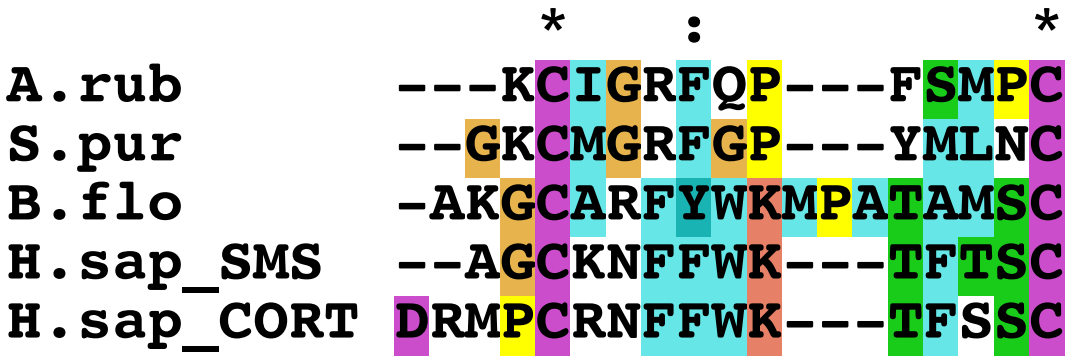


Fig. 2.13. *A. rubens* SMS-like precursor (ArSMSLP). (A) Sequence of ArSMSLP. An N-terminal signal peptide is represented in blue, a putative SMS-type neuropeptide (with cysteine (C) residues underlined) is represented in red and a dibasic cleavage site is represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative SMS-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* SMS-like precursor; *B.flo*, *Branchiostoma floridae* SMS-like precursor [JGI: 72051]; *H.sap_CORT*, *Homo sapiens* cortistatin (CORT) precursor [GI: 110645815]; *H.sap_SMS*, *Homo sapiens* SMS precursor [GI: 21619155]; *S.pur*, *Strongylocentrotus purpuratus* SMS-like precursor [GI: 390344260].

2.3.2.12. MCH-like precursor (ArMCHLP)

The *A. rubens* melanin-concentrating hormone (MCH)-like precursor (ArMCHLP) is an 88-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide and a putative MCH-type peptide with the predicted polypeptide sequence DRPNRREVTYCMDWIHNTWRPCRGRKAG (residues 61-88) termed ArMCH bounded by a dibasic cleavage site (KR) (**Fig. 2.14A; Fig. S39**). However, there is the possibility that ArMCH is a polypeptide sequence formed between residues 66-84 due to the presence of a putative dibasic cleavage site (RR) at residues 65 and 66 (EVTYCMDWIHNTWRPCRGRKAG) whilst the presence of a putative dibasic cleavage site (RK) at residues 85 and 86 may give rise to a version of both peptides in which the presence of a C-terminal glycine (G) residue on ArMCH is indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide (EVTYCMDWIHNTWRPCRGRKA-NH_2). It is noteworthy that this putative MCH-type peptide contains two cysteine (C) residues, which may form an intramolecular disulphide bridge in accordance with other MCH-type peptides (Lebl et al., 1988). ArMCHLP was identified through sequence similarity with Spnp14 in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012). However, upon the discovery of ArMCHLP and alignment with MCH-like precursors (**Fig. 2.14B**), it is likely that Spnp14 is a MCH-like precursor comprising a putative MCH-type neuropeptide with the predicted polypeptide sequence SRSGRKLRFCMDVIRNTWRLCRNTRSN, bounded by dibasic cleavage sites (KR).

MCH-type peptides are a family of neuropeptides with a phylogenetic distribution attributed to be indicative of a chordate origin (Jekely, 2013; Kawauchi, 2006; Kawauchi and Baker, 2004; Mirabeau and Joly, 2013). MCH was first identified in the chum salmon *Oncorhynchus keta* on account of its physiological

function of inducing body colour change (Kawauchi et al., 1983; Kawazoe et al., 1987). MCH-type peptides have subsequently been identified throughout the vertebrates (Nahon, 1994). However, to date, MCH-type peptides have not been identified outside of the vertebrates although MCH-like receptors have been identified in the cephalochordate *B. floridae* (Jekely, 2013) and in the hemichordate *S. kowalevskii*, in which there has been a large expansion (Mirabeau and Joly, 2013).

It is important to consider ArMCHLP in relation to MCH-like precursors/peptides identified in other phyla (**Fig. 2.14B**). ArMCHLP is the first member of the MCH-like neuropeptide family to be identified outside of the vertebrates. However, it is likely that Spnp14 previously identified in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) is also MCH-like based on the discovery of ArMCHLP and further support from neuropeptide alignments. Vertebrate MCH-type peptides have a characteristic cyclic structure due to the presence of two cysteine (C) residues that form an intramolecular disulphide bridge (Lebl et al., 1988). ArMCH and the MCH-type peptide derived from Spnp14 also have two cysteine (C) residues that are indicative of an intramolecular disulphide bridge. However, unlike vertebrate MCH-type peptides, ArMCH also has a C-terminal glycine (G) residue indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide.

To date, the physiological function of MCH-type peptides outside of the vertebrates and specifically in the echinoderms is currently unknown. To this extent, investigation would be facilitated by identification of receptors that mediate the effects of MCH-type peptides. Importantly, two candidate MCH-like receptors have been identified in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of MCH-type peptides in an echinoderm species (**Fig. A6J-K; Fig. S40-41**).

(A) Melanin-concentrating hormone (MCH)-like precursor (ArMCHLP)

MRPHVVLFACLAVPCLLLAQIVSCAPVYDGNNQIQPFDTDDWTGGDALGDQDFMMET
EKRDRPNRREVTYCMDWIHNTWRPCRGRKAG

(B) MCH-type neuropeptide multiple sequence alignments

```
                    .    :   *:      ..::  *
A.rub          DRPNRREVTYCMDWIHNTWRPCRGRKAG
S.pur          SRSGRK-LRFCMDVIRNTWRLCRNTRSN
S.kow          -R-GRKVLQVCIDWKKKTWKWCYFDL--
T.rub          -NADLDILLRCM--VGRVYRPCWEVE--
H.sap          ---DFD-MLRCM--LGRVYRPCWQV---
```

Fig. 2.14. ***A. rubens*** **MCH-like precursor (ArMCHLP). (A)** Sequence of ArMCHLP. An N-terminal signal peptide is represented in blue, a putative MCH-type neuropeptide (with cysteine (C) residues underlined) is represented in red and a dibasic cleavage site is represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative MCH-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* MCH-like precursor; *H.sap*, *Homo sapiens* MCH precursor [GI: 187445]; *S.kow*, *Saccoglossus kowalevskii* MCH-like precursor [GI: 187231810]; *S.pur*, *Strongylocentrotus purpuratus* MCH-like precursor [GI: 109402760]; *T.rub*, *Takifugu rubripes* MCH precursor [GI: 410918650].

2.3.2.13. CT-like precursor (ArCTLPP)

The *A. rubens* calcitonin (CT)-like precursor (ArCTLPP) is a 114-residue precursor protein comprising a predicted 21-residue N-terminal signal peptide and a 40-residue putative CT-type neuropeptide with the predicted polypeptide sequence NGESRGCSGFGGCGVLTIGHNAAMRMLAESNSPFGASGPG (residues 61-100) termed ArCT bounded by dibasic cleavage sites (KR) (**Fig. 2.15A; Fig. S42**). The presence of a C-terminal glycine (G) residue is indicative of post-translational modification giving rise to a C-terminal amide group on the mature polypeptide (NGESRGCSGFGGCGVLTIGHNAAMRMLAESNSPFGASGP-NH_2). It is noteworthy that this putative CT-type neuropeptide contains two cysteine (C) residues (residues 67 and 73), which are likely to form a disulphide bridge in accordance with CT/CT-gene related peptide (CGRP)-type peptides in other phyla.

CT was first discovered in mammals as a peptide that is released from parafollicular cells of the thyroid gland and inhibits intestinal calcium ion (Ca^{2+}) absorption and inhibits osteoclast activity in bones (Wendelaar Bonga and Pang, 1991). CT is encoded by a gene that also encodes CGRP, with alternative splicing giving rise to either prepro-CT (exons 1, 2, 3 and 4) or prepro-CGRP (exons 1, 2, 3, 5 and 6) (Amara et al., 1982). Thus, whilst the CT/CGRP gene encodes both CT and CGRP, mRNA transcripts encoding both peptides in tandem are not generated (Strand, 1999).

CT-type peptides have subsequently been identified in other vertebrates (Ogoshi et al., 2006) and deuterostomian invertebrates including urochordates (Sekiguchi et al., 2009), the cephalochordate *B. floridae* (Mirabeau and Joly, 2013), the hemichordate *S. kowalevskii* (Mirabeau and Joly, 2013) and the echinoderm species *S. purpuratus* (Rowe and Elphick, 2012) and *A. japonicus* (Rowe et al., 2014). In the protostomes, the first CT-type peptide to be discovered was diuretic

hormone (DH31) in the cockroach *Diploptera punctata* (Furuya et al., 2000). Subsequently, DH31-type peptides have been identified in other ecdysozoans including a number of arthropod species (Furuya et al., 2000) and nematodes (Mirabeau and Joly, 2013) and also in the lophotrochozoans (Mirabeau and Joly, 2013). Thus, CT/DH31-type peptides are a family of neuropeptides with a widespread phylogenetic distribution indicative of an origin that dates back to the common ancestor of the Bilateria (Jekely, 2013; Mirabeau and Joly, 2013).

It is important to consider ArCTLPP in relation to CT/DH31-like precursors/peptides identified in other phyla (**Fig. 2.15B**). ArCT contains a C-terminal amidated proline (P) motif, which appears to be a conserved feature of CT-type peptides, with the exception of CGRP. It would therefore appear that the C-terminal region of invertebrate CT-type peptides are more CT-like than CGRP-like. ArCT also has two cysteine (C) residues, which are likely to form a disulphide bridge in the N-terminal of the putative CT-type peptide. This feature is comparable to CT/CGRP-type peptides identified throughout the vertebrates (Ogoshi et al., 2006). It should be noted, however, that CT/CGRP-type peptides identified in other invertebrate species do not share this feature. For example, DH31-type peptides identified in the arthropods do not have these two cysteine (C) residues (Furuya et al., 2000). Identification of a CT-type peptide in the urochordate *C. intestinalis* with two cysteine (C) residues indicated at the time of discovery that this feature may be a unique characteristic of CT/CGRP-type peptides in chordates (Sekiguchi et al., 2009). However, there are two genes encoding CT-type peptides in the annelid *Capitella teleta*, one comprising a CT-type peptide with two N-terminal cysteine (C) residues and one comprising a DH31-type peptide without two N-terminal cysteine (C) residues (Mirabeau and Joly, 2013; Veenstra, 2011). It therefore appears that there may have been an ancient duplication of a CT-like gene, with one copy

retaining the N-terminal cysteine (C) residues and the other copy losing the N-terminal cysteine (C) residues in the protostomian lineage before subsequent loss of the CT-type peptide and retention of the DH31-type peptide in the arthropod lineage (Rowe et al., 2014).

It is also important to consider ArCTLPP in relation to CT-like precursors/peptides identified in other echinoderm species (**Fig. 2.15B**). ArCTLPP is the third member of the CT-like neuropeptide family to be identified in an echinoderm after the identification of Spnp4 (SpCTLPP) in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) and Ajnp4 (AjCTLPP) in the sea cucumber *A. japonicus* (Rowe et al., 2014). ArCTLPP, along with SpCTLPP, contains a single CT-type peptide. Interestingly, AjCTLPP is the first CT-like precursor that comprises two CT-type peptides in tandem (Rowe et al., 2014). However, a CT-like precursor comprising two CT-type peptides has also been identified in *S. kowalevskii* (Freeman et al., 2008; Mirabeau and Joly, 2013; Rowe et al., 2014) suggesting that independent intragenic duplication events may have given rise to genes encoding two CT-type peptides in both the vertebrate lineage, the hemichordate lineage and in the holothurian lineage.

To date, the physiological function of CT-type peptides in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of CT-type peptides. In this regard, a candidate CT-like receptor has been identified in the sea urchin *S. purpuratus* (Burke et al., 2006) but not in the sea cucumber *A. japonicus* (Rowe et al., 2014). Importantly, a candidate CT-like receptor has also been identified in the starfish *A. rubens* and may facilitate investigation into the physiological functions of CT-type peptides in an echinoderm species (**Fig. A4A; Fig. S43**).

(A) Calcitonin (CT)-like precursor (ArCTLPP)

MKPTTVLTLAVFCTLYTIITAASISRDDDMFDVTGDDLRQLAKKVDTYARNNEIQSL
LKRNGESRGCSGFGGCGVLTIGHNAAMRMLAESNSPFGASGPGKRRRSVDAVANQEA

(B) CT-type neuropeptide multiple sequence alignments

```
                                  *                                 *
A.rub          NGESRGCS-GFGG-CGVLTIGHNAAMRML-----AESNS-PFGASGPG
S.pur          ---SKGCG-SFSG-CMQMEVAKNRVAALL-----RNSNAHLFGLNGPG
S.kow          GRGTSACG-GFAT-CKQLEYGR--KYATS-----KADLS-HFGATSPG
B.flo1         --GKIACK---TAWCMNNRLSH--NLSSL-----DNPTD--TGVGAPG
B.flo2         -----DCS---TLTCFNQKLAH--ELAMD-----NQRTDTANPYS-PG
B.flo3         -----KCE---SGTCVQMHLADRLRLGLG-----HNMFTNTGPES-PG
B.flo4         ------NN---ESTCQTYLLKV---RKFT-----REFIHHAAGPGEPG
H.sap_Calc     ------CG--NLSTCMLGTYTQ--DFNKF-----HTFPQTAIGVGAPG
H.sap_CGRP1    -----ACD---TATCVTHRLAG--LLSRSGGVVKNNFVPTNVGSKAFG
C.tel          -----TCQFNLGGHCATESAAS--VADHW-----HYLNS---PLS-PG
L.gig1         -----SCSLRLGGMCLTENLNA--AANQY-----EYLSS---GLS-PG
L.gig2         -----SCNLNLGFHCQTDEYSS--IADMY-----DFLQS---ALS-PG
```

Fig. 2.15. ***A. rubens*** **CT-like precursor (ArCTLPP). (A)** Sequence of ArCTLPP. An N-terminal signal peptide is represented in blue, a putative CT-type neuropeptide (with cysteine (C) residues underlined) is represented in red, a C-terminal glycine (G) residue is represented in orange and dibasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative CT-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* CT-like precursor; *B.flo1*, *Branchiostoma floridae* CT-like precursor 1 [GI: 260826569]; *B.flo2*, *Branchiostoma floridae* CT-like precursor 2 [GI: 260826567]; *B.flo3*, *Branchiostoma floridae* CT-like precursor 3 [GI: 260826573]; *Bflo4*, *Branchiostoma floridae* CT-like precursor 4 [GI: 260812099]; *H.sap_Calc*, *Homo sapiens* CT precursor [GI: 179819]; *H.sap_CGRP1*, *Homo sapiens* CT-related polypeptide α [GI: 269784661]; *C.tel*, *Capitella teleta* CT-like precursor [GI: 161220966]; *L.gig1*, *Lottia gigantea* CT-like precursor 1 [GI: 163526287]; *L.gig2*, *Lottia gigantea* CT-like precursor 2 [GI: 676481265]; *S.kow*, *Saccoglossus kowalevskii* CT-like precursor [GI: 187217193]; *S.pur*, *Strongylocentrotus purpuratus* CT-like precursor [GI: 115767208].

2.3.2.14. CRH-like precursor (ArCRHLP)

The *A. rubens* corticotropin-releasing hormone (CRH)-like precursor (ArCRHLP) is a 130-residue precursor protein comprising a predicted 28-residue N-terminal signal peptide and a putative CRH-type peptide with the predicted polypeptide sequence QGLSVSPIFPIQRIRLNAIERDRQDQVDQAEANQGLFQIAG (residues 87-127) termed ArCRH bounded by dibasic cleavage sites (KR/RK) (**Fig. 2.16A; Fig. S44**). The presence of an N-terminal glutamine (Q) residue and a C-terminal glycine (G) residue is indicative of post-translational modifications giving rise to an N-terminal pyro-glutaminyl (pQ) residue and a C-terminal amide group on the mature peptide.

CRH-type peptides are a family of neuropeptides with a phylogenetic distribution indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). CRH is a hypothalamic neurohormone that stimulates release of adrenocorticotropic hormone (ACTH) in response to stress in mammals (Lovejoy, 2005). CRH-type peptides have been identified in other vertebrates (Lovejoy and Balment, 1999) and deuterostomian invertebrates including urochordates (Lovejoy and Barsyte-Lovejoy, 2010), the cephalochordate *B. floridae* (Mirabeau and Joly, 2013) and the hemichordate *S. kowalevskii* (Jekely, 2013; Mirabeau and Joly, 2013). However, whilst it has been proposed that CRH-type peptides have been identified in the echinoderm species *S. purpuratus* (Jekely, 2013; Mirabeau and Joly, 2013) and *A. japonicus* (Rowe et al., 2014), ArCRHLP is likely to be the first true precursor comprising a CRH-type peptide to be identified in an echinoderm species.

CRH-type peptides have been found to be orthologous to diuretic hormone DH44 identified in insects and to egg-laying hormone (ELH) identified in molluscs (Jekely, 2013; Lovejoy and Jahan, 2006; Mirabeau and Joly, 2013). DH44 was first identified in the fruit fly *D. melanogaster* and has been implicated in stimulating fluid secretion by Malpighian tubules through elevated cyclic adenosine

monophosphate (cAMP) levels (Cabrero et al., 2002). DH44-type peptides have subsequently been identified in a number of insect species but have yet to be identified in nematodes (Jekely, 2013; Mirabeau and Joly, 2013). ELH was first characterised in the mollusc *A. californica* (Scheller et al., 1983) and has subsequently been implicated in coordinating the process of egg-laying in *Aplysia* species (Conn and Kaczmarek, 1989). ELH-type peptides have also been identified in other molluscan species and in the annelids, where evidence of roles in promoting gamete release has been obtained (Salzet et al., 1997; Vreugdenhil et al., 1988).

It is important to consider ArCRHLP in relation to CRH-like precursors that have previously been identified in other echinoderms and CRH/DH44/EHL-like precursors identified in other phyla (**Fig. 2.16B**). CRH-like precursors containing two putative CRH-type peptides separated by an acidic spacer have recently been identified in the sea urchin *S. purpuratus* (Jekely, 2013; Mirabeau and Joly, 2013) and the sea cucumber *A. japonicus* (Rowe et al., 2014). It had been hypothesised that the occurrence of two CRH-type peptides in the echinoderms was likely a consequence of intragenic duplication in the echinoderm lineage with CRH/DH44-like precursors in other phyla containing only one CRH-type peptide (Rowe et al., 2014). However, upon the discovery of ArCRHLP, it appears that both of the CRH-like precursors identified in the echinoderm species *S. purpuratus* (Jekely, 2013) and *A. japonicus* (Rowe et al., 2014) are in fact PDF-like precursors based on neuropeptide alignments (see **chapter 2.3.2.15**). Further support of this notion is the presence of two PDF-type peptides in these echinoderm precursors, which is consistent with the presence of two PDF-type peptides in PDF-like precursors identified in other phyla. In accordance with CRH/DH44-like precursors identified in other phyla, ArCRHLP has only one CRH-type peptide, which opposes the hypothesis of an intragenic duplication in the echinoderm lineage (Rowe et al.,

2014).

To date, the physiological function of CRH-type peptides in the echinoderms is currently unknown. However, it has been proposed that CRH-type peptides may be involved in regulating the release of TSH-type hormones in the echinoderms by virtue of non-mammalian vertebrate CRH triggering the release of TSH (Rowe et al., 2014). To this extent, investigation would be facilitated by identification of receptors that mediate the effects of CRH-type peptides. In this regard, a candidate CRH-like receptor orthologous to mammalian CRHR1 and CRHR2 (Lovejoy, 2005) has been identified in the sea urchin *S. purpuratus* (Mirabeau and Joly, 2013). Similarly, an orthologue of mammalian CRHR1 and CRHR2 has also been identified in the sea cucumber *A. japonicus* (Rowe et al., 2014). Importantly, two candidate CRH-like receptors have been identified in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of CRH-type peptides in an echinoderm species (**Fig. A4B-C; Fig. S45-46**).

(A) Corticotropin-releasing hormone (CRH)-like precursor (ArCRHLP)

MNDLQRLILLVSLGTFALLLCLPACTEAQPLGLFKFEYDDLLDPSFEADDPRNPRRL
SRQQILRRLNDLAMSRSGSGPGYTIPRKRQGLSVSPIFPIQRIRLNAIERDRQDQVD
QAEANQGLFQIAGRKR

(B) CRH-type neuropeptide multiple sequence alignments

```
                                                *                                 :     *
A.rub       -----------------QGLSVSPIFPIQRIRLNAIERDRQDQ-----VDQAEANQGLFQIAG
S.kow       --------------NPTSGHSIGLSLDIIRDRLEKAESERMQQEKQSDINRYRQNTELMTGLG
B.flo       ---------------NPPHVAMDISTMLLGELLGAGGDGERAT--------TDHNIDLFRQAG
H.sap_CRH   --------------SEEPPISLDLTFHLLREVLEMARAEQLAQQ----AHSNRKLMEI---IG
C.tel1      ----------------SSPLSFGTDIDAISTMLASRRMGKHGR-------MRSVRHRLA-ALG
C.tel2      MMDLAKYGMPLKQAALQNDVRESLDDRILRGMLLRQNAQEFKTK----SRQGKVMQKLF-GIG
C.tel3      -------------EPTQAQMNVQLELHVLRNMLRSQERRPLSSD----PTEAKSMNKLF-GIG
C.tel4      --------------VPKPLMLVDITPDLLRAMLYRQQRQRMAAK----IRPQEPKSDLR-FIG
L.gig1      -----------------SRLSINQELKSLANLLVLRENKRRE------AQKTKLRSKLL-SIG
L.gig2      ----------------AGRLSINGALSSLADLLVS-ENQRRDR-----LESMELRQRLQ-YLG
```

Fig. 2.16. ***A. rubens*** **CRH-like precursor (ArCRHLP). (A)** Sequence of ArCRHLP. An N-terminal signal peptide is represented in blue, a putative CRH-type neuropeptide is represented in red, a C-terminal glycine (G) residue is represented in orange and dibasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative CRH-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* CRH-like precursor; *B.flo*, *Branchiostoma floridae* CRH-like precursor [GI: 260786674]; *C.tel1*, *Capitella teleta* CRH-like precursor 1 [GI: 161303031]; *C.tel2*, *Capitella teleta* CRH-like precursor 2 [JGI: 190906]; *C.tel3*, *Capitella teleta* CRH-like precursor 3 [JGI: 190906]; *C.tel4*, *Capitella teleta* CRH-like precursor 4 [JGI: 194553]; *H.sap*_CRH, *Homo sapiens* CRH precursor [GI: 30583744]; *L.gig1*, *Lottia gigantea* CRH-like precursor 1 [GI: 676493124]; *L.gig2*, *Lottia gigantea* CRH-like precursor 2 [GI: 163524672]; *S.kow*, *Saccoglossus kowalevskii* CRH-like precursor 1 [GI: 281433636].

2.3.2.15. PDF-like precursor (ArPDFLP)

The *A. rubens* pigment-dispersing factor (PDF)-like precursor (ArPDFLP) is a 104-residue precursor protein comprising a predicted 22-residue N-terminal signal peptide and two putative PDF-type peptides with the predicted polypeptide sequences LGDNDFFQATYNDAQARQRQRVLQSYLDDRMASVG (residues 31-65) termed ArPDF1 and NFDEDVYHQEGLDNEFVRRLMAKYFDGVA (residues 73-101) termed ArPDF2 bounded by dibasic cleavage sites (KR/RR) (**Fig. 2.17A; Fig. S47**). The presence of a C-terminal glycine (G) residue on ArPDF1 is indicative of a post-translational modification giving rise to a C-terminal amide group on the mature peptide (LGDNDFFQATYNDAQARQRQRVLQSYLDDRMASV-NH_2).

PDF-type peptides are a family of neuropeptides with a phylogenetic distribution recently attributed to be indicative of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). The insect PDF-type peptides are orthologous to the crustacean pigment-dispersing hormone (PDH) family (Rao and Riehm, 1993) and have been implicated in regulating circadian rhythms in insects (Bahn et al., 2009; Lee et al., 2009). PDF-type peptides have subsequently been identified in a number of ecdysozoans including other arthropods (Rao and Riehm, 1993) and nematodes (Janssen et al., 2009) and in lophotrochozoans including molluscs (Veenstra, 2010) and annelids (Veenstra, 2011). In the deuterostomes, a candidate PDF-type peptide has been identified in the hemichordate *S. kowalevskii* (Mirabeau and Joly, 2013). However, PDF-type peptides have not been identified in the vertebrates, the cephalochordate *B. floridae* or, until now, an echinoderm species. Moreover, PDF-like receptors have been identified in the ambulacrarians suggesting that the PDF-like neuropeptide signalling system can be traced back to the common ancestor of the bilaterians (Jekely, 2013; Mirabeau and Joly, 2013).

It is important to consider ArPDFLP in relation to PDF-like

precursors/peptides that have been identified in other phyla (**Fig. 2.17B**). It can be postulated that precursors previously identified in the sea urchin *S. purpuratus* (Jekely, 2013; Mirabeau and Joly, 2013) and in the sea cucumber *A. japonicus* (Rowe et al., 2014) as CRH-like are in fact PDF-like based on the discovery of ArCRHLP and further support from neuropeptide alignments (**Fig. 2.16B; Fig. 2.17B**). ArPDFLP has two PDF-type peptides – termed ArPDF1 and ArPDF2 – with ArPDF1 having a C-terminal glycine (G) residue indicative of C-terminal amidation. Similarly, PDF-like precursors in *S. purpuratus* and in the sea cucumber *A. japonicus* also have two PDF-type peptides. However, both of the PDF-type peptides in *S. purpuratus* (SpPDF1 and SpPDF2) have a C-terminal glycine (G) residue indicative of C-terminal amidation, whilst only one of the PDF-type peptides in *A. japonicus* (AjPDF2) has a C-terminal glycine (G) residue indicative of C-terminal amidation. Similarly, PDF-like precursors identified in other phyla also have two PDF-type peptides, with C-terminal amidation on one or both of these PDF-type peptides (Jekely, 2013; Mirabeau and Joly, 2013). Further support of the notion that ArPDFLP is orthologous to the *S. purpuratus* and *A. japonicus* PDF-like precursors previously thought to be CRH-like is the conservation of an N-terminal NDF motif common to ArPDF1, SpPDF1 and AjPDF1. Interestingly, ArPDF2 contains an N-terminal NFD motif suggesting that the presence of two PDF-type peptides could be due to an intragenic duplication. However, it is not clear whether this intragenic duplication occurred in the common ancestor of the bilaterians or whether this intragenic duplication occurred in the echinoderm lineage due to a noticeable difference between PDF-type peptide length in other phyla compared to a similar PDF-type peptide length in the echinoderms.

To date, the physiological function of PDF-type peptides in the deuterostomes and specifically in the echinoderms is currently unknown. To this

extent, investigation would be facilitated by identification of receptors that mediate the effects of PDF-type peptides. In this regard, a candidate PDF-like receptor has been identified in the sea urchin *S. purpuratus* (Jekely, 2013; Mirabeau and Joly, 2013). Importantly, two candidate PDF-like receptors have been identified in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of PDF-type peptides in an echinoderm species (**Fig. A5D-E; Fig. S48-49**).

(A) Pigment-dispersing factor (PDF)-like precursor (ArPDFLP)

MTQLTLLAVCGSVLLLVGLTHCTDEQREKRLGDNDFFQATYNDAQARQRQRVLQSYL
DDRMASVGKRDGLKRNFDEDVYHQEGLDNEFVRRLMAKYFDGVARRR

(B) PDF-type neuropeptide multiple sequence alignments

```
                                    :       .     :    :   ..
A.rub1      -LGDNDFFQATYND-AQARQRQRVLQSYLDDRMASVG
A.rub2      NFDEDVY--------HQEGLDNEFVRRLMAKYFDGVA
S.pur1      -IADNDFAAMRHQERSNSMRRTRLLQA-MNEMLAKAG
S.pur2      SLAQNDYMMVR-----QDLANGRLYRSLMDRMLSEAG
C.tel       ----------------NSGMLDAVIN--MP-DLFKAG
L.gig       ----------------NAGTVDHLLN--FP-DL-SVG
D.mel       ----------------NSELINSLLS--LPKNMNDAG
```

Fig. 2.17. *A. rubens* PDF-like precursor (ArPDFLP). (A) Sequence of ArPDFLP. An N-terminal signal peptide is represented in blue, putative PDF-type neuropeptides are represented in red, a C-terminal glycine (G) residue is represented in orange and dibasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative PDF-type neuropeptides derived from precursors of representative species across the Bilateria. *A.rub*, *Asterias rubens* PDF-like precursor; *C.tel*, *Capitella teleta* PDF-like precursor [JGI: 204689]; *D.mel*, *Drosophila melanogaster* PDF-like precursor [GI: 281362639]; *L.gig*, *Lottia gigantea* cerebrin precursor [GI: 676458325]; *S.pur*, *Strongylocentrotus purpuratus* PDF-like precursor [GI: 115899431].

2.3.2.16. PP/orcokinin-like precursor 1 (ArPPLNP1)

The *A. rubens* pedal peptide (PP)/orcokinin-like neuropeptide precursor (ArPPLNP1) is a partial 325-residue precursor protein comprising a predicted 31-residue N-terminal signal peptide and seven putative PP/orcokinin-type neuropeptides (**Fig. 2.18A; Fig. S50**). These include single copies of the polypeptide sequences GRTNMYGSSQLSRLSSGFN (residues 101-155) termed ArPPLNP1a, GRSSFAGSSRLTNLGSGFT (residues 142-160) termed ArPPLNP1b, GRSSFTGSSRLTNLASGFN (residues 176-194) termed ArPPLNP1c, GRSAFSGSRGLTNLASGFN (residues 211-229) termed ArPPLNP1d, GRSSFAGVSGLTHLGSGFN (residues 238-256) termed ArPPLNP1e, GRSAFSGSKGLTNLASGFN (residues 273-291) termed ArPPLNP1f and a single copy of the partial polypeptide sequence GRTSLSGSSGLTHLSSG (residues 309-325) termed ArPPLNP1g bounded by dibasic cleavage sites (KR/KK).

PP/orcokinin-type peptides are a family of neuropeptides with a phylogenetic distribution recently attributed to be indicative of an ancestral bilaterian origin (Rowe and Elphick, 2012). PP was first discovered in the mollusc *A. californica* as a peptide that causes contraction of pedal muscles (Hall and Lloyd, 1990; Lloyd and Connolly, 1989) and has subsequently been shown to cause beating of cilia associated with the pedal sole (Longley and Peterman, 2013). PP-type peptides were thought to be unique to the lophotrochozoans, with identification of PPs in molluscs and annelids (Conzelmann et al., 2011; Veenstra, 2011). However, the recent discovery of Spnp6 (SpPPLNP1) and Spnp7 (SpPPLNP2) in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) and Ajnp7 (AjPPLNP2) in the sea cucumber *A. japonicus* (Rowe et al., 2014) has shown that PPs are of an ancestral bilaterian origin. PPs have also subsequently been identified in the ecdysozoans including in the nematode *C. elegans* (Rowe and Elphick, 2012).

PPs have been found to be orthologous to orcokinin-type peptides identified in the arthropods (Rowe and Elphick, 2012). Orcokinin was first isolated from abdominal nerve cord extracts of the crayfish *Orconectus limosus* on account of its function of stimulating hindgut myoactivity (Stangier et al., 1992). Orcokinin-type peptides have subsequently been identified in a number of arthropods and attributed with a range of physiological functions including the stimulation of the prothoracic gland and the regulation of ecdysteroidogenesis in the silk moth *Bombyx mori* (Yamanaka et al., 2011) and regulation of circadian and seasonal physiology in the cockroach *Leucophaea maderae* (Hofer and Homberg, 2006; Soehler et al., 2011; Wei and Stengl, 2011). It would therefore appear that the PP/orcokinin-like neuropeptide signalling system could be traced back to the common ancestor of the bilaterians. Interestingly, however, PP/orcokinin-type peptides have not yet been identified in the vertebrates, suggesting that the PP/orcokinin-like neuropeptide signalling system may have been lost in the vertebrate lineage.

It is important to consider ArPPLNP1 in relation to PP/orcokinin-like precursors/peptides identified in other phyla (**Fig. 2.18B**). PP/orcokinin-like precursors in other phyla comprise multiple copies of PP/orcokinin-type peptides. Similarly, ArPPLNP1 contains seven putative PPs. However, unlike in other phyla, including in the mollusc *A. californica* comprising four PP-like precursors (Moroz et al., 2006), only one PP-like precursor has been identified thus far in the starfish *A. rubens*. There appears to be little sequence similarity between PPs derived from ArPPLNP1 and PP/orcokinin-type peptides in other phyla. However, each of the PPs derived from ArPPLNP1 excluding ArPPLN1g, which may be a partial sequence, include a C-terminal SGFx (x is a hydrophobic residue) motif common to PP in the mollusc *A. californica*. Moreover, PPs derived from ArPPLNP1 and PP from the mollusc *A. californica* has a characteristic distribution of hydrophobic and

hydrophilic residues throughout each of the respective peptides.

It is also important to consider ArPPLNP1 in relation to PP-like precursors/peptides identified in other echinoderm species (**Fig. 2.18B**). ArPPLNP1 is the fourth member of the PP-like neuropeptide family to be identified in an echinoderm after the identification of SpPPLNP1 and SpPPLNP2 in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) and AjPPLNP2 in the sea cucumber *A. japonicus* (Rowe et al., 2014). It appears that ArPPLNP1 shares the highest sequence similarity with SpPPLNP1 and SpPPLNP2, which also contain a C-terminal SGFx motif common to PP in the mollusc *A. californica*. Interestingly, however, each of the PPs derived from AjPPLNP2 do not have the C-terminal SGFx motif but have a C-terminal glycine (G) residue indicative of C-terminal amidation, which is atypical to the majority of PP/orcokinin-type peptides identified in other phyla to date. However, two PPs derived from SpPPLNP1 also have C-terminal glycine (G) residues indicative of C-terminal amidation.

To date, the physiological function of PP/orcokinin-type peptides in the echinoderms is currently unknown with preliminary pharmacological investigations of synthetic SpPPLN1c in the sea urchin species *E. esculentus* not revealing any myoactive effects on the contractile activity of tube foot or oesophagus preparations *in vitro* (Rowe and Elphick, 2012). Identification of candidate PP/orcokinin-like peptide receptors in the echinoderm species *S. purpuratus*, *A. japonicus* or *A. rubens* may facilitate investigation into the physiological functions of PP's in an echinoderm species. However, to date, no PP/orcokinin-like receptors have been identified or characterised throughout the bilaterians.

(A) Partial (3') pedal peptide (PP)/orcokinin-like precursor 1 (ArPPLNP1)

MARCGGETPRALRTWWLVGLFCLALSLVCQAEIEANDVELVPEPEETKANEILLEEL
RDELFNELLQELEDEVAKGLTPEGRDLFHKRLTSLNSEWRAKRGRTNMYGSSQLSRL
SSGFNKRTLTDDESALEDLLDDAEVKRGRSSFAGSSRLTNLGSGFTKKSDPGVWLDS
EDKRGRSSFTGSSRLTNLASGFNKRDEDAYLLDDFLSKRGRSAFSGSRGLTNLASGF
NKREEAVKRGRSSFAGVSGLTHLGSGFNKRGDFLEDVYANEDKRGRSAFSGSKGLTN
LASGFNKRSDGDQSLWEENDVKRGRTSLSGSSGLTHLSSG

(B) PP-type neuropeptide multiple sequence alignments

```
                                    .  .:
A.rub       SAFSGSRGLTNLASGFN------
S.pur1      -GF-NSGAMEPLGAGFF------
S.pur2      --F-GSGSLEPMSSGFY------
C.tel       -AFDQIGGLSNFG-GFD------
L.gig       -------PMDSLASGLVG-----
A.cal       -PLDSVYGTHGMS-GFA------
D.mel       ----NFDEIDKASASFSILNQLV
```

Fig. 2.18. ***A. rubens*** **partial (3') PP/orcokinin-like precursor (ArPPLNP1). (A)** Sequence of ArPPLNP1. An N-terminal signal peptide is represented in blue, putative PP-type neuropeptides are represented in red and dibasic cleavage sites are represented in green. **(B)** ClustalX v.2.1 multiple sequence alignment of putative PP-type neuropeptides derived from precursors of representative species across the Bilateria. *A.cal*, *Aplysia californica* PP precursor-1 [GI: 325297152]; *A.rub*, *Asterias rubens* PP/orcokinin-like precursor; *C.tel*, *Capitella teleta* orcokinin precursor [GI: 161190484]; *D.mel*, *Drosophila melanogaster* orcokinin precursor [GI: 442624594]; *L.gig*, *Lottia gigantea* orcokinin precursor [GI: 163527842]; *S.pur1*, *Strongylocentrotus purpuratus* PP/orcokinin-like precursor 1 [GI: 390335272]; *S.pur2*, *Strongylocentrotus purpuratus* PP/orcokinin-like precursor 2 [GI: 390352581].

2.3.3. Precursors of neuropeptides and peptide hormones that contain multiple cysteine (C) residues

Transcripts encoding neuropeptide and peptide hormone precursors that contain multiple cysteine (C) residues and disulphide bridges have been identified by analysis of the neural transcriptome of the starfish *A. rubens* (**Table 4**). These precursors will be discussed in groups or on a one-by-one basis (or in groups).

***A. rubens* precursor**	**Precursor family**	**Key peptide/precursor characteristics**
ArGPA2-1-2	GPA2-type	Subunit with ten conserved C residues
ArGPB5-1-3	GPB5-type	Subunit with multiple conserved C residues
ArBALPP/ ArBBLPP	Bursicon-like	Subunit(s) with eleven conserved C residues
ArGSSLP	GSS-like	A-chain with four conserved C residues; B-chain with two conserved C residues
ArRLP	Relaxin/INSL-like	A-chain with four conserved C residues; B-chain with two conserved C residues
ArIGF1-2	IGF-like	A-E domain structure; A-chain with four conserved C residues; B-chain with two conserved C residues

Table 4. Summary of key peptide/precursor characteristics of neuropeptides/peptide hormones with multiple cysteine (C) residues. *GPA2*, glycoprotein hormone α-2; *GPB5*, glycoprotein hormone β-5; *GSS*, gonad-stimulating substance; *IGF*, insulin-like growth factor; *INSL*, insulin. Note that the table is not exhaustive and additional similarities and differences between neuropeptides/precursors are expanded on in the respective sub-headings.

2.3.3.1. GPA2/5-type precursors (ArGPA2-1-2 and ArGPB5-1-3)

The *A. rubens* glycoprotein hormone α-2 (GPA2)-type precursor 1 (ArGPA2-1) is a 135-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide followed by a 111-residue polypeptide sequence sharing sequence similarity with GPA2 subunits (**Fig. 2.19A; Fig. S51**). A second GPA2-type precursor (ArGPA2-2) was identified as a 130-residue precursor protein comprising a predicted 28-residue N-terminal signal peptide followed by a 102-residue polypeptide sequence sharing sequence similarity with GPA2 subunits (**Fig. 2.19B; Fig. S52**). It is noteworthy that both ArGPA2-1 and ArGPA-2 contain ten cysteine (C) residues, which are likely to form five disulphide bridges in accordance with glycoprotein hormone subunits in other phyla (**Fig. 2.21**) (Lovejoy, 2005).

The *A. rubens* glycoprotein hormone β-5 (GPB5)-type precursor 1 (ArGPB5-1) is a 136-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide followed by a 112-residue polypeptide sequence sharing sequence similarity with GPB5 subunits (**Fig. 2.19C; Fig. S53**). A second GPB5-type precursor (ArGPB5-2) is a 141-residue precursor protein comprising a predicted 31-residue N-terminal signal peptide followed by a 110-residue polypeptide sequence sharing sequence similarity with GPB5 subunits (**Fig. 2.19D; Fig. S54**). A third GPB5-type precursor (ArGPB5-3) is a 130-residue precursor protein comprising a predicted 30-residue N-terminal signal peptide followed by a 100-residue polypeptide sequence sharing sequence similarity with GPB5 subunits (**Fig. 2.19E; Fig. S55**). It is noteworthy that ArGPB5-1 contains eight cysteine (C) residues whilst both ArGPB5-2 and ArGPB5-3 contain ten cysteine (C) residues, which are likely to form four or five disulphide bridges, respectively, in accordance with glycoprotein hormone subunits in other phyla (**Fig. 2.21**) (Lovejoy, 2005).

The glycoprotein hormones are a family of cysteine (C)-rich polypeptide

hormones with a phylogenetic distribution indicative of an evolutionary ancestry that dates back to the common ancestor the Bilateria (Park et al., 2005). The prototypical members of the glycoprotein hormone family are the mammalian gonadotropins LH, FSH, choriogonadotropin (CG) and TSH (Pierce and Parsons, 1981). The gonadotropins are heterodimeric glycoproteins comprising a common α-subunit and a unique β-subunit, which is crucial for receptor specificity. The α-subunit has ten cysteine (C) residues forming five disulphide bridges whilst the β-subunit has twelve cysteine (C) residues forming six disulphide bridges, with the formation of α/β dimers necessary for biological activity (Lovejoy, 2005). The α-subunit is common to the α/β dimer within a species whilst the β-subunit determines hormone specificity (Fiddes and Goodman, 1979; Fiddes and Goodman, 1980).

Sequencing of the human genome revealed a novel member of the glycoprotein hormone family member, which is now known as thyrostimulin (Hsu et al., 2002). Thyrostimulin is a heterodimeric glycoprotein comprising an α-subunit termed GPA2 and a β-subunit termed GPB5 and it acts as a ligand for TSH receptors (Nakabayashi et al., 2002). GPA2 and GPB5-type subunits have subsequently been identified in a number of vertebrates (Hsu et al., 2002) and in deuterostomian invertebrates including urochordates (Campbell et al., 2004; Dos Santos et al., 2009), the cephalochordate *B. floridae* (Holland et al., 2008), the hemichordate *S. kowalevskii* (Dos Santos et al., 2009) and the echinoderm species *S. purpuratus* (Burke et al., 2006) and *A. japonicus* (Rowe et al., 2014). GPA2 and GPB5-type subunits have also been identified in ecdysozoans including arthropods (Sudo et al., 2005) and nematodes (Park et al., 2005) and in lophotrochozoans including molluscs (Veenstra, 2010) and annelids (Veenstra, 2011). Thus, the phylogenetic distribution of GPA2 and GPB5-type subunits is indicative of these subunits being of an ancestral bilaterian origin, with the α- and β-subunits of vertebrate LH, FSH, CG and

TSH evolving from GPA2 and GPB5-type subunits, respectively, in the vertebrate lineage (Sower et al., 2009).

To date, the physiological function of GPA2 and GPB5-type subunits in invertebrate species is not particularly well understood although it has been hypothesised that these peptide hormones may modulate ion and osmotic balance in insects (Sellami et al., 2011). The physiological function of GPA2 and GPB5-type subunits in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of GPA2 and GPB5-type subunits. In this regard, a candidate receptor orthologous to DLGR1 identified in *D. melanogaster* (Sudo et al., 2005) has been identified in the sea urchin *S. purpuratus*. Importantly, a candidate receptor has also been identified in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of GPA2 and GPB5-type subunits in an echinoderm species (**Fig. A7A; Fig. S56**).

(A) Glycoprotein hormone α-2-type precursor 1 (ArGPA2-1)

MQWCKVVLTLMSLSVCLCSLFATAQQRPYWERSGCHLVGYSTLVQIPGCHKTRVDMN
ACRGYCVTYTLLSTFNQIVSNNIRYSSRGTCCAIGDTHDVIVILACENNEQKSVTYK
SAASCSCTLCTQEDASQLNNV

(B) Glycoprotein hormone α-2-type precursor 2 (ArGPA2-2)

MTMKVSVTFIYVACTAALLILVSSPVKGAWEPTAGCHLVGYRKEVRVPGCHIEYVKM
NACRGYCMTYSFLSDTATLERSGGTQLFTSHGSCCSITSTHDVHITLQCENNQVYKD
TFKSAKTCSCALCSTQ

(C) Glycoprotein hormone β-5-type precursor 1 (ArGPB5-1)

MDQGAAYTSIVMVTLVMMWACALAINPVTTTNCYVHTAMKHLVEKPGCRPHELVVFG
CWGRCDTNEVPSLDPPFVEAYHPVCTLTNYEDVKVKLPDCDPEVDPTYTYQSALSCG
CANIDDSSTKYSYRPDYFVSEK

(D) Glycoprotein hormone β-5-type precursor 2 (ArGPB5-2)

MFFKCQARHRWTATFLPLFALALLIGTIVEGARYQKQSMPFLDVSCRVREYTKYEAK
LPGCMDEVVPARGCYGRCQSFEVPVLLPPHKASSHKMCLVEEIELRSVELSDCLPGV
NRTFVYQSAVRCRCKKCIESNTFCARN

(E) Glycoprotein hormone β-5-type precursor 3 (ArGPB5-3)

MMATRRLTAVQNLAIFGFVVLVLSATLCTSQGLSCLPRQYIKYDAVKPGCRTQRITI
YGCFGRCHTSEIPKLLPPYKESNHAMCSYGQTESRVILLDDCDPGVDPTFQYEDALS
CACKKCEPWNTFCQGF

Fig. 2.19. *A. rubens* GPA2-type and GPB5-type precursors. (A) Glycoprotein hormone α-2-type precursor 1 (ArGPA2-1). **(B)** GPA2-type precursor 2 (ArGPA2-2). **(C)** Glycoprotein hormone β-5-type precursor 1 (ArGPB5-1) **(D)** GPB5-type precursor 2 (ArGPB5-2). **(E)** GPB5-type precursor 3 (ArGPB5-3). N-terminal signal peptides are represented in blue and putative GPA2-type and GPB5-type subunits are represented in red (with cysteine (C) residues underlined).

2.3.3.2. Bursicon α/β-like precursors (ArBALP and ArBBLP)

The *A. rubens* bursicon-α-like hormone precursor (ArBALP) is a 139-residue precursor protein comprising a predicted 35-residue N-terminal signal peptide followed by a 104-residue polypeptide sequence comprising eleven cysteine (C) residues, sharing sequence similarity with bursicon-α type subunits (**Fig. 2.20A; Fig. S57**). The *A. rubens* bursicon-β-like hormone precursor (ArBBLP) is a 142-residue precursor protein comprising a predicted 30-residue N-terminal signal peptide followed by a 112-residue polypeptide sequence comprising eleven cysteine (C) residues, sharing sequence similarity with bursicon-β-type subunits (**Fig. 2.20B; Fig. S58**). It is noteworthy that both ArBALP and ArBBLP contain eleven cysteine (C) residues, sharing sequence similarity with bursicon-type subunits (**Fig. 2.21**).

The bursicons, belonging to the glycoprotein hormone protein family (see **chapter 2.3.3.1**), are a family of polypeptide hormones with a phylogenetic distribution indicative of an ancestral bilaterian origin. Bursicon was originally discovered in insects on account of its effect in cuticular tanning (Fraenkel et al., 1966) and has subsequently been shown to be a heterodimeric protein formed by the subunits bursicon-α and bursicon-β encoded by separate precursors (Luo et al., 2005). Bursicon-type subunits have subsequently been identified in the ecdysozoans including arthropods (Van Loy et al., 2007) and have been shown to function in cuticle hardening and ecdysis in crustacean species (Chung et al., 2012; Webster et al., 2013; Wilcockson and Webster, 2008). Bursicon subunits have also been identified in the lophotrochozoans including molluscs (Veenstra, 2010) and annelids (Veenstra, 2011) and have also recently been identified in a number of deuterostomian invertebrates including the echinoderm species *S. purpuratus* (Burke et al., 2006), *A. japonicus* (Rowe et al., 2014) and now *A. rubens*, indicative of the bursicon-type peptide hormone system being of an ancestral bilaterian origin.

To date, the physiological function of bursicons in non-arthropod species is not particularly well understood and in the echinoderms is currently unknown. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of bursicons. In this regard, two candidate bursicon-like receptors orthologous to bursicon-like receptors in *Drosophila* and other insects (Bai and Palli, 2010; Baker and Truman, 2002) and leucine-rich repeat containing GPCRs in vertebrates (LGR4-6) (Deng et al., 2013) have been identified in the sea urchin *S. purpuratus*; however, to date, no bursicon-like receptors have been identified in the sea cucumber *A. japonicus* (Rowe et al., 2014). Importantly, a candidate bursicon-like receptor has been identified in the starfish *A. rubens* and this may facilitate investigation into the physiological functions of bursicons in an echinoderm species (**Fig. A7B; Fig. S59**).

(A) Bursicon α-like precursor (ArBALP)

MDSHHRNTGTASATVVVKCALIAILLLGVISQCSAVCRRQPLIHSIVHEGCQTKRLR
TFGCRGTCNSYSRVSPTDYTQMERSCQCCQESQHVVGFVELNCPSLSPPTQIVEFRH
VRSCSCRPCNSVSGVPRVTRLEDLE

(B) Bursicon β-like precursor (ArBBLP)

MAPMQHHHHLAAIFIFSVLSMCLLPDLVQAVRRGPAGTCEVGISFITVEEEFESSDG
GSVISCTGTTTVNQCEGECVTTSTPSVTEPTGYSKICKCCREQSLRPKQVMLSDCYD
SAGNAITGQQYPVYVPEPASCSCQKCSR

Fig. 2.20. *A. rubens* bursicon-type precursors. (A) Bursicon α-like precursor (ArBALP). **(B)** Bursicon β-like precursor (ArBBLP). N-terminal signal peptides are represented in blue and putative bursicon-type subunits are represented in red (with cysteine (C) residues underlined).

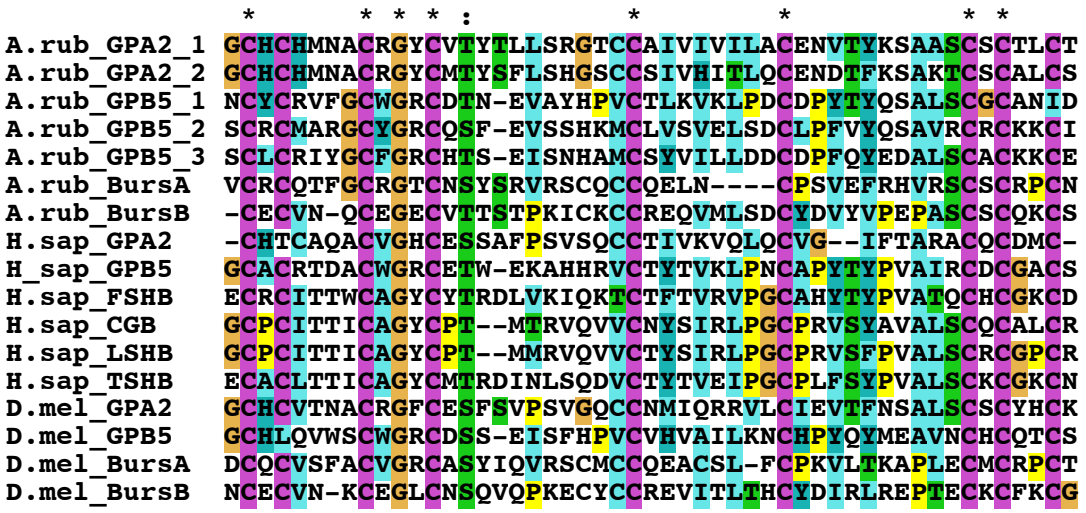


Fig. 2.21. Conservation of bursicon-like, GPA2-type and GPB5-type subunits in representative phyla and species across the animal kingdom. ClustalX v.2.1 multiple sequence alignment of putative bursicon-like, GPA2-type and GPB5-type subunits derived from precursors of representative species across the Bilateria. *A.rub_GPA2_1*, *Asterias rubens* GPA2-type precursor 1; *A.rub_GPA2_2*, *Asterias rubens* GPA2-type precursor 2; *A.rub_GPB5_1*, *Asterias rubens* GPB5-type precursor 1; *A.rub_GPB5_2*, *Asterias rubens* GPB5-type precursor 2; *A.rub_GPB5_3*, *Asterias rubens* GPB5-type precursor 3; *A.rub_BursA*, *Asterias rubens* bursicon α-type precursor; *A.rub_BursB*, *Asterias rubens* bursicon β-type precursor; *D.mel_GPA2*, *Drosophila melanogaster* GPA2 precursor [GI: 320546230]; *D.mel_TSHB*, *Drosophila melanogaster* glycoprotein hormone β-subunit-related precursor [GI: 21427595]; *D.mel_BursA*, *Drosophila melanogaster* bursicon-α precursor [GI: 665394724]; *D.mel_BursB*, *Drosophila melanogaster* bursicon-β precursor [GI: 62392020]; *H.sap_GPA2*, *Homo sapiens* GPA2 precursor [GI: 189491650]; *H.sap_GPB5*, *Homo sapiens* GPB5 precursor [GI: 21427593]; *H.sap_FSHB*, *Homo sapiens* follicle-stimulating hormone (FSH) β-precursor [GI: 124014246]; *H.sap_CGB7*, *Homo sapiens* chorionic gonadotropin (CG) β-polypeptide-7 precursor [GI: 15451749]; *H.sap_LSHB*, *Homo sapiens* luteinising hormone (LH) β-precursor [GI: 15431286]; *H.sap_TSHB*, *Homo sapiens* thyroid stimulating hormone (TSH) β-precursor [GI: 47479817].

2.3.3.3. GSS-like precursor (ArGSSLP)

The *A. rubens* gonad-stimulating substance (GSS)-like precursor (ArGSSLP) is a 109-residue precursor protein comprising a predicted 26-residue N-terminal signal peptide and two putative gonad-stimulating substance (GSS)-like peptides (**Fig. 2.22; Fig. S60**). These include the polypeptide sequences AEKYCDEDFHMAVYRTCTEH (residues 27-46) termed the Ar B-chain and PETYVGMGSYCCLVGCTRDQLSQVC (residues 85-109) termed the Ar A-chain separated by a connecting peptide (C-peptide) (residues 49-82). It is noteworthy that the Ar B-chain contains two cysteine (C) residues (residues 31 and 43) and the Ar A-chain contains four cysteine (C) residues (residues 95, 96, 100 and 109), which are likely to form disulphide bridges and a peptide heterodimer between both chains, based on previous structural characterisation of GSS in the starfish species *A. pectinifera* (Mita et al., 2009).

GSS was first discovered on account of inducing gamete release in the starfish *A. rubens* (Chaet and McConnaughy, 1959) and was identified as a peptide hormone (Kanatani et al., 1971; Kanatani and Shirai, 1970) before subsequent purification in the starfish species *A. pectinifera* (Mita et al., 2009). GSS has been shown to induce oocyte maturation and ovulation and is thus functionally analogous the vertebrate gonadotropin LH (Mita et al., 2009). However, subsequently, GSS has been identified as a member of the insulin/insulin-like growth factor (IGF)/relaxin superfamily (Mita, 2013a). The insulin/IGF/relaxin superfamily is divided into two distinct sub-classes: the insulin/IGF sub-class and the relaxin/insulin-like (INSL) sub-class, with GSS belonging to the relaxin/INSL sub-class (Mita et al., 2009).

It is important to consider ArGSSLP in relation to the GSS precursor identified in *A. pectinifera* (Mita et al., 2009). ArGSSLP and the GSS precursor in *A. pectinifera* both have an A-chain and a B-chain linked by a connecting C-peptide

organised in the same manner as in relaxin/INSL-like peptide hormones. The A-chains both comprise the cysteine motif CCxxxCxxxxxxxxC, characteristic of the insulin/IGF/relaxin superfamily. However, the final residue in the A-chain is a cysteine (C) residue, indicative of the relaxin/INSL sub-class as opposed to the insulin/IGF sub-class, which have an asparagine (N) residue after this residue (Mita, 2013a). Similarly, the B-chains both have the cysteine motif CxxxxxxxxxxC, which is also characteristic of the insulin/IGF/relaxin superfamily. It could therefore be predicted that the A-chain and B-chain of ArGSSLP would form a peptide heterodimer through disulphide bonds, consistent with the structural properties of GSS in *A. pectinifera* (Mita et al., 2009).

Gonad-stimulating substance (GSS)-like precursor (ArGSSLP)

MANYRLILEATCLLVLLINTALYAEAAEKYCDEDFHMAVYRTCTEHKRSGRSAFSLN
DFFRSNSKRTAGSPRPDDDFFLTMQKRPETYVGMGSYCCLVGCTRDQLSQVC

Fig. 2.22. *A. rubens* GSS-like precursor (ArGSSLP). An N-terminal signal peptide is represented in blue, putative GSS-like neuropeptides (with cysteine (C) residues underlined) are represented red, a putative C-peptide is represented in black and dibasic cleavage sites are represented in green.

2.3.3.4. Relaxin-like precursor (ArRLP)

The *A. rubens* relaxin-like precursor (ArRLP) is a 119-residue precursor protein comprising a predicted 30-residue N-terminal signal peptide and two putative relaxin-like peptides (**Fig. 2.23; Fig. S61**). These include the polypeptide sequences RSDHASVKHFCGLEFSYAVVTACGEA (residues 31-56) termed the Ar B-chain and QDYQGMATYCCTNGCTISQLTNSGIC (residues 94-119) termed the Ar A-chain separated by a C-peptide (residues 59-91). The presence of an N-terminal glutamine (Q) residue in the Ar B-chain is indicative of a post-translational modification giving rise to an N-terminal pyro-glutaminyl (pQ) residue (pQDYQGMATYCCTNGCTISQLTNSGIC). It is noteworthy that the Ar B-chain contains two cysteine (C) residues (residues 41 and 53) and the Ar A-chain contains four cysteine (C) residues (residues 103, 104, 108 and 119), which are likely to form disulphide bridges and a peptide heterodimer between both chains based on the presence of this feature in the insulin and the relaxin-like peptide hormone family (Wilkinson et al., 2005).

It is important to consider ArRLP in relation to insulin/IGF/relaxin superfamily precursors identified in other echinoderm species. ArRLP is only the third insulin/IGF/relaxin superfamily precursor to be identified in an echinoderm after the identification of SpIGF1 and SpIGF2 in the sea urchin *S. purpuratus* (Burke et al., 2006). SpIGF1 has characteristics of the both the insulin/IGF and relaxin/INSL sub-classes, with the presence of dibasic cleavage sites at the N-terminal and C-terminal of the A-chain and B-chain indicative of cleavage of the C-peptide (Burke et al., 2006). However, the presence of a C-chain and D-chain is indicative of the insulin/IGF-subclass (Burke et al., 2006). SpIGF2 also has characteristics of the insulin/IGF sub-class, with no dibasic cleavage sites present at the N-terminal and C-terminal of the A-chain and B-chain indicative of a lack of cleavage of the C-peptide

(Burke et al., 2006; Lapraz et al., 2006). However, unlike SpIGF1 and SpIGF2, it appears that ArRLP has characteristics of the relaxin/INSL sub-class with the presence of dibasic cleavage sites at the N-terminal and C-terminal of the A-chain and B-chain indicative of cleavage of the C-peptide (Burke et al., 2006). Moreover, ArRLP does not contain a D-domain and E-domain indicative of the insulin/IGF sub-class. However, similar to GSS, the relaxin/INSL-like peptide derived from ArRLP does not have the relaxin-specific receptor-binding motif RxxxRxxI/V characteristic of the vertebrate relaxin-like peptides (Bullesbach and Schwabe, 2000).

Interestingly, four orthologous relaxin/INSL genes are present in teleost fish and mammals (Good-Avila et al., 2009; Wilkinson and Bathgate, 2007). It has therefore been suggested that these vertebrate relaxin/INSL genes are a result of whole genome duplication (WGD) during the evolution of vertebrates from an invertebrate ancestor (Hoffmann and Opazo, 2011; Yegorov and Good, 2012). More specifically, it has been suggested that the relaxin/INSL genes are products of an ancestral system that originally consisted of three genes (two of which trace their origins back to the invertebrates) before subsequent diversification of the system through WGD (Yegorov and Good, 2012). Thus, it can be hypothesised that both ArGSSLP and ArRLP represent the ancient invertebrate orthologues of the vertebrate relaxin/INSL family.

To date, the physiological function of GSS has been well studied in *A. pectinifera* but not much is known about the function of relaxin/INSL-like peptide hormones in the echinoderms. To this extent, investigation would be facilitated by the identification of receptors that mediate the effects of GSS and relaxin/INSL-like peptide hormones. However, despite some understanding of the GSS signal transduction pathway (Mita, 2013a; Mita, 2013b), the identity of the GSS receptor remains unknown. It can be hypothesised that GSS and relaxin/INSL-like peptide

hormones bind to relaxin-like receptors. However, the A-chains and B-chains derived from both ArGSSLP and ArRLP do not contain the vertebrate relaxin-specific receptor-binding motif RxxxRxxI/V (Bullesbach and Schwabe, 2000). Further studies would therefore be necessary to identify the receptors that mediate the effects of GSS and relaxin/INSL-like peptide hormones in starfish and other echinoderm species.

Relaxin-like precursor (ArRLP)

MTSCSHQMLALLSAVYILIFFLGGLPAVHARSDHASVKHFCGLEFSYAVVTACGEAK
RSIRSAPFFDMFPVFKSPERIPADFDDSSMIHVRKRQDYQGMATYCCTNGCTISQLT
NSGIC

Fig. 2.23. *A. rubens* relaxin-like precursor (ArRLP). An N-terminal signal peptide is represented in blue, putative relaxin-like neuropeptides (with cysteine (C) residues underlined) are represented in red, a putative C-peptide is represented in black and dibasic cleavage sites are represented in green.

2.3.3.5. IGF-like precursors (ArIGF1-2)

The *A. rubens* insulin-like growth factor 1 precursor (ArIGF1) is a 355-residue precursor protein comprising a predicted 27-residue N-terminal signal peptide and characteristic polypeptide A-E domains found in IGF-like precursors (**Fig. 2.24A; Fig. S62**) (Reinecke and Collet, 1998). It is noteworthy that the B-domain contains two cysteine (C) residues (residues 32 and 44) and the A-domain contains four cysteine (C) residues (residues 91, 92, 96 and 105), which are likely to form disulphide bridges and a peptide heterodimer between both domains based on the presence of this feature in the relaxin/INSL-like peptide hormone family. ArIGF1 has dibasic cleavage sites at the C-terminal and N-terminal of the B-domain (residues 28-55) and A-domain (residues 83-106) respectively, which would allow for the removal of the intervening C-peptide (residues 58-80) as in the relaxin/INSL-like peptide hormone family. However, the ArIGF1 A-domain extends to the D-domain (residues 107-133) and E-domain (residues 134-355), a feature found in IGF-like precursors (Reinecke and Collet, 1998).

The *A. rubens* IGF 2 precursor (ArIGF2) is a 343-residue precursor protein comprising a predicted 22-residue N-terminal signal peptide and characteristic polypeptide A-E domains found in IGF-like precursors (with no D-domain detected) (**Fig. 2.24B; Fig. S63**) (Reinecke and Collet, 1998). It is noteworthy that the B-domain contains two cysteine (C) residues (residues 41 and 53) and the A-domain contains four cysteine (C) residues (residues 66, 67, 71 and 80), which are likely to form disulphide bridges and a peptide heterodimer between both domains based on the presence of this feature in the relaxin/INSL-like peptide hormone family. However, unlike vertebrate insulin and relaxin, ArIGF2 does not contain dibasic cleavage sites at the C-terminal and N-terminal of the B-domain (residues 23-46) and A-domain (residues 61-81) respectively, indicating that the intervening C-peptide

(residues 47-80) is not processed for removal. Moreover, the ArIGF2 A-domain extends to the E-domain (residues 82-343), a feature also found in IGF-like precursors (Reinecke and Collet, 1998).

It is important to consider ArIGF1 and ArIGF2 in relation to IGF-like precursors identified and subsequently characterised in the sea urchin *S. purpuratus* (Burke et al., 2006; Perillo and Arnone, 2014). ArIGF1 and SpIGF1 both have an A-E domain organisation as found in IGF-like precursors and are more relaxin/INSL-like in relation to potential removal of the C-peptide. ArIGF2 and SpIGF2 both have a similar A-E domain organisation (with no D-domain detected) and are more IGF-like in relation to the C-peptide (Reinecke and Collet, 1998). Both the ArIGF1/2 A-domains have the cysteine motif CCxxxCxxxxxxxxC, characteristic of the insulin/IGF/relaxin superfamily. Similarly, the B-domains both have the cysteine motif CxxxxxxxxxxxC, characteristic of the insulin/IGF/relaxin superfamily. Importantly, two disulphide bridges connecting the B-domain and A-domain and a single intra-disulphide bridge associated to the A-domain have been identified in SpIGF1 and SpIGF2 (Perillo and Arnone, 2014) suggesting that this is a feature of ArIGF1 and ArIGF2. However, it has been suggested that the lack of conservation in the C-E domains suggests that these domains are not involved in a conserved function and that their role may be species-specific (Perillo and Arnone, 2014).

To date, the physiological function of IGF-like peptides in non-chordate deuterostomes and in particular echinoderms has not been extensively studied. However, it has previously been shown that mammalian insulin stimulates growth of the sea urchin embryo and that IGFs may be expressed in the sea urchin larval gut (de Pablo et al., 1988) and adult starfish gut (Wilson and Falkmer, 1965) suggesting a role in the feeding process. Moreover, microarray data confirm expression in the sea urchin embryo (Samanta et al., 2006). The identification of SpIGF1 and SpIGF2

in the sea urchin *S. purpuratus* has made it possible to investigate the roles of IGFs in the echinoderms (Burke et al., 2006). It has recently been shown that SpIGF1 is expressed in both the stomach and intestine of feeding larva with differential expression dependent on nutrient availability, suggesting a role for SpIGF1 in the feeding process (Perillo and Arnone, 2014). Moreover, it has been shown that SpIGF2 is expressed in the gastrula foregut, whilst a putative IGF-like receptor (SpInsr) is expressed in the mesodermal cells at the tip of the archenteron (Perillo and Arnone, 2014). Taken together, this suggests a role for SpIGF2 as a growth signal to stimulate coelomic pouch development (Perillo and Arnone, 2014). The discovery of ArIGF1 and ArIGF2 will facilitate investigation into the physiological functions of IGF-like peptides in a starfish species for the first time.

(A) Insulin-like growth factor 1-like precursor (ArIGF1)

MFRNTSTMRALLLLDVIFVALVLPITAWPKICGEQLVETVSLVCSTRGFYSHRDSKR
DVEVFQNERAAKSFLGSRIGSRQRRRTGRIATECCDRICSFDIVESYCNPWPVAIES
RDPPLSPVAPGRVREDKSADVDYMYNPDVVDVEEANSVIQREEDLIDDIETQEQEIE
QDEEQNMQTLPEEDAEDTDIREPEDVEESFPVPVPTKKRRKVEGRRSKESKNKGGKS
EGKNKKRSGSREGGRSSRRSRGKSSRSKKQRDGRERSKRWEGLDTSHPVKEPTARSV
LGRVDTRPFRNFLYNRYTVDEKRDTERESYRAVAPLTGYNSHRGGSQPDNHPTLAAL
YNLAVKLAKGLQH

(B) Insulin-like growth factor 2-like precursor (ArIGF2)

MNQYQLIVLFEVLAHASMLNYASPVQLCGRELTETLRSICGDRGYYSPGQSFSRRAP
THDGIATRCCQSLCESSILETYCNLPAPPSQTQPSTAAPTTTTKMAPLTEDRRTKDV
VVDYSDQLATEGSQMSRVDGVLTHDTVTNRSKTTTESNEGSYDNEEGAPYDKPDDSS
PSERGESIQDEDNEVNKPEPNNIRDNSKERGRNRTHKGVSSERRANNSRRRGLSSER
RGSSSSRREEKLRRRRQRHRERELREQRKQSNSKRKSKGDKKDHSVAATTPLAVQER
PLKNGGRNSTSGEHSSVNGTETDTAGAGSPEVKKDDLITTITAVLSDMIGFQPDNGN
R

Fig. 2.24. ***A. rubens*** **IGF-like precursors. (A)** Insulin-like growth factor 1-like precursor (ArIGF1). (**B)** IGF 2-like precursor (ArIGF2). N-terminal signal peptides are represented in blue, putative B-domains and A-domains (with cysteine (C) residues underlined) are represented in red, putative C-peptides are represented in black, a putative D-domain is represented in olive, putative E-domains are represented in maroon and dibasic cleavage sites are represented in green.

Fig. 2.25. Conservation of insulin/IGF/relaxin-like peptide hormones in representative phyla and species across the Bilateria. ClustalX v.2.1 multiple sequence alignment of putative insulin/IGF/relaxin-like peptide hormones derived from precursors of representative species across the Bilateria. *A.rub_IGF1*, *Asterias rubens* insulin-like growth factor (IGF1) precursor 1; *A.rub_IGF2*, *Asterias rubens* IGF precursor 2; *A.rub_GSS*, *Asterias rubens* GSS precursor; *A.rub_Rel*, *Asterias rubens* relaxin-like precursor; *D.mel_INSL1*, *Drosophila melanogaster* insulin-like (INSL) peptide precursor 1 [GI: 317423340]; *D.mel_INSL2*, *Drosophila melanogaster* INSL peptide precursor 2 [GI: 442631434]; *D.mel_INSL3*, *Drosophila melanogaster* INSL peptide precursor 3 [GI: 221331056]; *D.mel_INSL4*, *Drosophila melanogaster* INSL peptide precursor 4 [GI: 442631435]; *D.mel_INSL5*, *Drosophila melanogaster* INSL peptide precursor 5 [GI: 320545737]; *D.mel_INSL6*, *Drosophila melanogaster* INSL peptide precursor 6 [GI: 442614930]; *H.sap_INS*, *Homo sapiens* insulin precursor [GI: 186429]; *H.sap_IGF1*, *Homo sapiens* IGF precursor [GI: 49456676]; *H.sap_Rel1*, *Homo sapiens* relaxin-1 precursor [GI: 35932]; *H.sap_Rel2*, *Homo sapiens* relaxin-2 precursor [GI: 35926]; *H.sap_Rel3*, *Homo sapiens* relaxin-3 precursor [GI: 17484095].

2.3.4. Precursors of putative neuropeptides that appear to be unrelated to known bilaterian neuropeptide families

Neuropeptide precursors previously identified in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) that are not known to be of a bilaterian origin were used for BLAST analysis of the neural transcriptome of the starfish *A. rubens*. Homologues of five neuropeptide precursors previously identified in *S. purpuratus* are discussed below (**Table 5**). However, homologues of nine sea urchin neuropeptide precursors (Spnp6-9, Spnp10, Spnp13, Spnp16-17 and Spnp20) were not identified in *A. rubens*. It could be possible that these neuropeptide precursors are unique to *S. purpuratus*. However, identification of Ajnp7, Ajnp9 and Ajnp16 in the sea cucumber *A. japonicus* (Rowe et al., 2014) suggests that it is likely that homologues for at least some of these neuropeptide precursors exist in other echinoderms but may be too divergent to be found by BLAST analysis.

***A. rubens* precursor**	**Precursor family**	**Key peptide/precursor characteristics**
ArANPP	AN-type	Multiple peptides with an N-terminal AN motif
Arnp11	11-type	Single/multiple peptide(s) with six conserved C residues
Arnp15a-b	15-type	Single/multiple peptide(s) with six conserved C residues
Arnp18	18-type	Single/multiple peptide(s) with eight conserved C residues

Table 5. Summary of key peptide/precursor characteristics of non-bilaterian neuropeptide precursor families. Note that the table is not exhaustive and additional similarities and differences between neuropeptides/precursors are expanded on in the respective sub-headings.

2.3.4.1. AN peptide-type precursor (ArANPP)

The. *A. rubens* AN peptide-type precursor (ArANPP) is a 274-residue precursor protein comprising a predicted 24-residue N-terminal signal peptide and six putative neuropeptides with an N-terminal dipeptide alanine-asparagine (AN) motif (**Fig 2.26; Fig. S64**). These include two copies of the polypeptide sequence ANYHARGGKPRGGF (residues 179-192 and 219-232) termed ArANP1 and single copies of the polypeptide sequences ANHYASGRQRTRGKAGY (residues 56-72) termed ArANP2, ANYRASVSRGG (residues 95-105) termed ArANP3, ANYRATMNGG (residues 134-143) termed ArANP4 and ANYHASRGRQRGNG (residues 258-271) termed ArANP5 bounded by dibasic cleavage sites (KR/RR). The presence of a C-terminal glycine (G) residue on ArANP3-5 are indicative of post-translational modifications giving rise to C-terminal amide groups on the mature peptides (e.g. ANYRASVSRG-NH_2) based on the presence of C-terminal amidation in SpANP1, SpANP2 and SpANP3 derived from Spnp5 (SpANPP) in the sea urchin *S. purpuratus* (Menschaert et al., 2010; Rowe and Elphick, 2012).

The AN-peptides are a family of peptides that to date have only been identified in the echinoderms (Rowe et al., 2014; Rowe and Elphick, 2012). The AN-peptides do not appear to share any sequence similarities with neuropeptides identified in other phyla. It has previously been hypothesised that the presence of a putative amidated C-terminal lysine (K) residue in SpANP1, SpANP3 and SpANP5 and the presence of a putative amidated C-terminal arginine (R) residue in AjANP2 derived from Ajnp5 (AjANPP) in the sea cucumber *A. japonicus* suggests that amidated basic residues may be important structural features of the AN-peptides (Rowe et al., 2014). Interestingly, the feature of basic residue amidation is indicative of potential relationships with neuropeptides that have been discovered in the TK family of neuropeptides. However, the identification of a putative TK-like precursor

in *A. rubens* (see **chapter 2.3.2.6**) means that AN-type peptides are unlikely to be candidate members of the TK-like neuropeptide family and these similarities are likely due to convergence.

However, despite the relationship of AN-type peptides with peptides in other phyla being currently unclear, it is important to consider ArANPP in relation to AN-type precursors/peptides identified in other echinoderm species. ArANPP is the third member of the AN-type neuropeptide family to be identified after the identification of Spnp5 (SpANPP) in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) and Ajnp5 (AjANPP) in the sea cucumber *A. japonicus* (Rowe et al., 2014). ArANPP, along with SpANPP and AjANPP, contain multiple copies of putative peptides with an N-terminal dipeptide AN motif. ArANPP comprises six putative AN-type peptides, SpANPP comprises thirteen putative AN-type peptides whilst AjANPP comprises only three putative AN-type peptides. However, this could be in part due to the AjANPP sequence being a partial sequence (Rowe et al., 2014). It should also be noted that these putative AN-type peptides could be amidated or non-amidated at the C-terminal; ArANPP contains putative AN-type peptides both with and without a C-terminal glycine (G) residue as a substrate for amidation. However, AN-type peptides derived from both SpANPP and AjANPP all contain a C-terminal glycine (G) residue as substrates for amidation, with the exception of SpANP5 (Rowe and Elphick, 2012). Importantly, C-terminal amidation has been shown for SpANP1, SpANP2 and SpANP3 derived from SpANPP (Menschaert et al., 2010; Rowe and Elphick, 2012).

To date, the physiological functions of the AN-type peptides in the echinoderms is currently unknown; preliminary pharmacological investigations of synthetic SpANP2 on *in vitro* tube foot and oesophagus preparations from the sea urchin species *E. esculentus* have not revealed any myoactive effects of the peptide

(Rowe and Elphick, 2012).

AN peptide-type precursor (ArANPP)

MARLDLILLLLSVVVAAVVLQCRADDENTDGQGAADVDEVLKQLYAEDENDDDKRAN
HYASGRQRTRGKAGYRRVTKSDMDAVAGLEDMDEEKRANYRASVSRGGKKGFRRYTK
SDGDEEEIPELEEMDEEKRANYRATMNGGRKPHRFNRFTKADGDGEEEEGPEDLEGE
DIDEDKRANYHARGGKPRGGFRRYTKSDEDQDAMEQAPAEEMDEDKRANYHARGGKP
RGGFRRYTKSDEDLAMGEAAADELDEEKRANYHASRGRQRGNGRRN

Fig. 2.26. *A. rubens* AN peptide-type precursor (ArANPP). An N-terminal signal peptide is represented in blue, putative AN-type neuropeptides are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green.

2.3.4.2. Arnp11

Arnp11 is a 103-residue precursor protein comprising a 21-residue N-terminal signal peptide followed by an 82-residue polypeptide sequence (residues 22-103) that contains a dibasic cleavage site (RR) at residues 45 and 46 (**Fig. 2.27; Fig. S65**). The N-terminal region of the protein (residues 22-44) contains six acidic amino acid residues, which indicates that this part of the protein may function as an acidic spacer peptide. It can be proposed that it is the 57-residue polypeptide formed by residues 47-103 that may be a secreted bioactive neuropeptide. However, there is the possibility that the secreted bioactive neuropeptide is a 54-residue polypeptide formed between residues 50-103 due to the presence of a putative dibasic cleavage site (RK) at residues 48 and 49. It is noteworthy that the 57-residue polypeptide includes six cysteine (C) residues located at positions 54, 58, 61, 72, 76 and 97, suggesting the presence of up to three potential disulphide bridges that would confer tertiary structure on the polypeptide. However, alternatively, a homodimeric protein could be formed by up to six intermolecular disulphide bridges. It should also be noted, however, that the 57-residue polypeptide also includes potential monobasic

cleavage sites (R/K), so there remains the possibility it is cleaved into two or more bioactive neuropeptides.

Arnp11 was identified through sequence similarity with Spnp11 in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012). Spnp11 comprises a putative 54-residue polypeptide, which shares high sequence similarity with the putative 57-residue polypeptide in Arnp11, including the presence of six cysteine (C) residues. However, the putative bioactive neuropeptides derived from both Arnp11 and Spnp11 do not share any apparent sequence similarity with neuropeptides or peptide hormones identified in any other phyla. However, neuropeptides of a similar size and with six cysteine (C) residues have been identified in other animals. For example, moult-inhibiting hormone (MIH) is a 78-residue neuropeptide in the crustacean *Carcinus maenas* with six cysteine (C) residues that form three intramolecular disulphide bridges (Webster, 1991).

Arnp11

MRTALVTFLAMLLVGDLIVSALPIDNEQDSDPIFDHLYTRNIVERRSRKDLTKCISE
CVSCAKYAGLYADKCVRGCSSKTSGKGIINKTEFDAWSACEQFLHR

Fig. 2.27. *A. rubens* Arnp11 precursor. An N-terminal signal peptide is represented in blue, a putative neuropeptide (with cysteine (C) residues underlined) is represented in red and a dibasic cleavage site is green.

2.3.4.3. Arnp15a-b

Arnp15a is a 111-residue protein comprising a predicted 19-residue N-terminal signal peptide followed by a 92-residue polypeptide sequence (residues 20-111) that contains a dibasic cleavage site (KR) at residues 54 and 55 (**Fig. 2.28A; Fig. S66**). The N-terminal region of the protein (residues 12-53) contains eight acidic amino acid residues, indicating that this part of the protein may function as an acidic spacer peptide. It can be proposed that is the 53-residue polypeptide formed by residues 56-108 that may be a secreted bioactive neuropeptide. The presence of a C-terminal glycine (G) residue is indicative of a post-translational modification, giving rise to a C-terminal amide group on the mature peptide. However, the N-terminal region of the protein (residues 12-53) contains a potential dibasic cleavage (KR) site at residues 39 and 40, so there remains the possibility that the precursor forms two or more secreted bioactive neuropeptides. The presence of six cysteine (C) residues in the 53-residue polypeptide suggests that there may be up to three intramolecular disulphide bridges. Alternatively, a homodimeric protein could be formed by up to six intermolecular disulphide bridges.

Arnp15b is a 111-residue protein comprising a predicted 25-residue N-terminal signal peptide followed by an 86-residue polypeptide sequence (residues 26-111) that contains a dibasic cleavage site (KR) at residues 63 and 64 (**Fig. 2.28B; Fig. S67**). The N-terminal region of the protein (residues 26-62) contains six acidic amino acid residues, indicating that this part of the protein may function as an acidic spacer peptide. It can be proposed that is the 47-residue polypeptide formed by residues 65-111 that may be a secreted bioactive neuropeptide. However, the N-terminal region of the protein (residues 26-62) contains a potential dibasic cleavage site (KR) at residues 39 and 40, so there remains the possibility that the precursor forms two or more secreted bioactive neuropeptides. The presence of six cysteine (C)

residues in the 47-residue polypeptide suggests that there may be up to three intramolecular disulphide bridges. Alternatively a homodimeric protein could be formed by up to six intermolecular disulphide bridges.

Arnp15a-b were identified through sequence similarity with Spnp15 in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012). Spnp15 contains a putative 40-residue bioactive neuropeptide, which shares high sequence similarity with the putative bioactive neuropeptides in Arnp15a-b, including the presence of six cysteine (C) residues. However, these peptides derived from both Arnp15a-b and Spnp15 do not exhibit any apparent sequence similarity with neuropeptides identified in any other phyla. However, neuropeptides of a similar size and with six cysteine (C) residues have been identified in other animals including the 28-residue neuropeptide trissin in the fruit fly *D. melanogaster* with six cysteine (C) residues that form three intramolecular disulphide bridges (Ida et al., 2011).

(A) Arnp15a

MKSILCICAIALLVEVVIGAPFAEMAGDMSDEDSQLFSKRVRDIISPLSHDLDKRKA
NLCAMDCFSCFKMIRNVSPDQCVTGCQKKSISDGSYSYDRMWNRCSSYLTGRRR

(B) Arnp15b

MSASTITTLLLLSLAALSVFLPNASAEWSPDNDNNNDHKRAQMSAHDWLNSLLAGNDV
GHAKTKRIAKACAIDCLNCGMMFQVQAYSQGDCLTACQNDDNHRDPSCHEHITM

Fig. 2.28. *A. rubens* Arnp15a-b precursors. (A) Arnp15a. **(B)** Arnp15b. N-terminal signal peptides are represented in blue, putative neuropeptides (with cysteine (C) residues underlined) are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic/tribasic cleavage sites are represented in green.

2.3.4.4. Arnp18

Arnp18 is a 113-residue precursor protein comprising a predicted 27-residue N-terminal signal peptide followed by an 86-residue polypeptide sequence (residues 28-113) that contains a dibasic cleavage sites (KR) at residues 36 and 37 (**Fig. 2.29; Fig. S68**). It can be proposed that it is the 76-residue polypeptide formed by residues 38-113 that may be a secreted bioactive neuropeptide. It is noteworthy that this putative 76-residue polypeptide contains eight cysteine (C) residues, which may form up to four intramolecular disulphide bridges. Alternatively, a homodimeric protein could be formed by up to eight intermolecular disulphide bridges.

Arnp18 was identified through sequence similarity with Spnp18 in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012). Spnp18 contains a putative 70-residue bioactive neuropeptide, which shares high sequence similarity with the putative 76-residue bioactive neuropeptide in Arnp18, including the presence of eight cysteine (C) residues. However, these peptides derived from both Arnp18 and Spnp18 do not share any apparent sequence similarity with neuropeptides or peptide hormones identified in any other phyla. However, neuropeptides of a similar size and with eight cysteine (C) residues have been identified in other animals including the anti-gonadotropic peptide schistosomin in the pond snail *L. stagnalis* thought to contain four intramolecular disulphide bridges (Hordijk et al., 1991).

Arnp18

MQSSTVVTLATCLIIATIVTESASALAFQGSQDRAKRLFWVDKKSVNEDTTVCVRAS
SADDIAECFITECIKHQINCEMICGSDSEPCHSLCKAKKSDCAINCLEREQQSMSQ

Fig. 2.29. *A. rubens* Arnp18 precursor. An N-terminal signal peptide is represented in blue, a putative neuropeptide (with cysteine (C) residues underlined) is represented in red and a dibasic cleavage site is represented in green.

2.4. General discussion

Herein, the identification of a repertoire of novel neuropeptide precursors and a number of candidate receptors in the starfish *A. rubens* has provided a foundation to investigate the physiological functions of neuropeptide signalling systems in an echinoderm species, in many cases, for the very first time. Utilisation of this transcriptome dataset has been highlighted by the identification of the VP/OT-type peptide asterotocin and the NG peptide NGFFYamide, for which the physiological roles of these peptides has been investigated (see **chapters 3 and 4**). However, before undertaking further functional studies on other putative neuropeptides, it will be important to confirm these sequences through cloning and sequencing and potential post-translational modifications through the use of mass spectrometry (see **chapters 3 and 4**).

The identification of novel neuropeptide precursors in the starfish *A. rubens* has also provided important insights into the evolutionary origins of a number of neuropeptide families. As an echinoderm species, *A. rubens* is a deuterostomian invertebrate and therefore occupies an intermediate position in animal phylogeny, bridging model protostomian invertebrates (e.g. *Drosophila* and *C. elegans*) to vertebrates to provide important insights into neuropeptide diversity and its functional significance across the animal kingdom. Amongst these novel neuropeptide precursors, and perhaps most interestingly from an evolutionary perspective, precursors comprising two KP-type peptides (the first identified in a non-chordate species), a MCH-type peptide (the first identified outside of the vertebrates) and two TK-type peptides (the first identified in an ambulacrarian species) have been discovered. Thus, the identification of neuropeptide precursors in *A. rubens* has provided important new insights into the evolutionarily ancient origins of neuropeptide signalling systems.

It is important, however, to note that the list of neuropeptide precursors identified here in the starfish *A. rubens* is unlikely to be exhaustive. Neuropeptides are typically comprised of relatively short stretches of amino acids with relatively few conserved residues, with the number and length of repetitive bioactive neuropeptides often changing through evolution. Therefore, neuropeptide precursors can often be too divergent to be identified through BLAST analysis alone. For example, it would be reasonable to expect to find homologues of the neuropeptide precursors Spnp7/Ajnp7, Spnp9/Ajnp9 and Spnp17/Ajnp17 in *A. rubens*, which have been identified in the sea urchin *S. purpuratus* (Rowe and Elphick, 2012) and the sea cucumber *A. japonicus* (Rowe et al., 2014) respectively.

However, attempts at providing a more comprehensive insight into the complement of neuropeptide precursors in a transcriptome dataset are being provided. For example, a number of novel neuropeptide precursors have been identified in the starfish *A. rubens* using the neuropeptide prediction tool termed NPsearch (https://rubygems.org/gems/NpSearch), which utilises common characteristics of neuropeptides and their precursors to identify novel sequences computationally. The development of computational tools that allow the identification of novel neuropeptide precursors will be vital in providing a comprehensive list of sequences present in a given transcriptome dataset.

The starfish species *A. rubens* has been used as a model system for neuropeptide research for over twenty-five years, pioneered by the discovery of the first neuropeptides to be identified in an echinoderm – the SALMFamides S1 and S2 (Elphick et al., 1991a; Elphick et al., 1991b). Now, in a post-genomic and transcriptomic era, the generation of a neural transcriptome dataset has revitalised the use of *A. rubens* as a model system for neuropeptide research in the echinoderms and has begun to provide important insights into neuropeptide function and evolution

(see **chapters 3 and 4**) and will provide the basis to provide further insights into a number of other novel neuropeptide signalling systems.

3. Discovery and functional characterisation of a novel vasopressin/oxytocin (VP/OT)-type neuropeptide ("asterotocin") in the starfish *A. rubens*

3.1. Introduction

Vasopressin (VP) and oxytocin (OT) are highly related and evolutionarily conserved neurohypophyseal peptides with a plethora of known physiological functions in mammals and other vertebrates. VP (CYFQNCPRG-NH_2) has been classically viewed as regulating vasoconstriction and osmoregulation (Henderson and Byron, 2007; McCormick and Bradshaw, 2006), whilst OT (CYIQNCPLG-NH_2) has been viewed as a reproductive hormone involved in uterine contraction and the stimulation of lactation (Gimpl and Fahrenholz, 2001).

VP/OT-type peptides are also found outside of the vertebrates and have a widespread phylogenetic distribution indicative of an ancestry that dates back to at least as far as the common ancestor of the Bilateria (Jekely, 2013; Mirabeau and Joly, 2013; Stafflinger et al., 2008; van Kesteren et al., 1995). Thus, VP/OT-type peptides have been identified in vertebrates (Hoyle, 1999; Urano et al., 1992), protostomian invertebrates (Beets et al., 2012; Cruz et al., 1987; Garrison et al., 2012; Jekely, 2013; Mirabeau and Joly, 2013; Proux et al., 1987; Reich, 1992; van Kesteren et al., 1992) and recently in a number of deuterostomian invertebrates (Elphick and Rowe, 2009; Jekely, 2013; Kawada et al., 2008; Mirabeau and Joly, 2013; Ukena et al., 2008).

VP/OT-type peptides are highly conserved across the Bilateria (**Fig. 3.1**). All known members of the VP/OT neuropeptide family have two conserved cysteine (C) residues in positions 1 and 6 of the mature neuropeptide that have been shown to form a disulphide bridge and confer a cyclic conformation that is important for biological activity (Hruby et al., 1990; Sawyer, 1977). VP/OT-type peptides have

hydrophobic residues in positions 2 and 3 of the mature neuropeptide, whilst other conserved features include a proline (P) residue in position 7 (exceptions include the urochordate *Ciona intestinalis* and the hemichordate *Saccoglossus kowalevskii*) and a C-terminal glycine (G) residue in position 9, which is usually a substrate for C-terminal amidation (exceptions include the urochordate *C. intestinalis* and the nematode *Caenorhabditis elegans*).

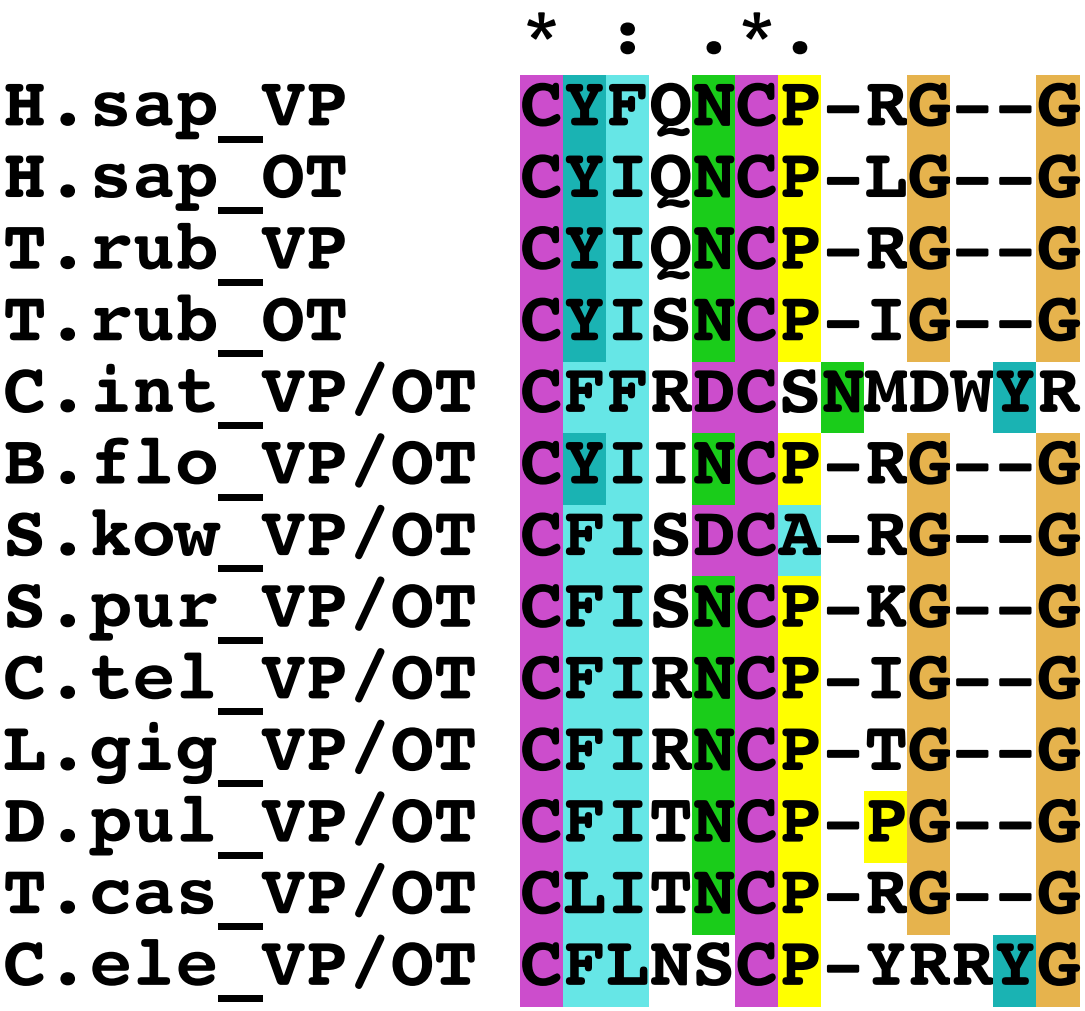


Fig. 3.1. Conservation of VP/OT-type peptides in representative phyla and species across the Bilateria. ClustalX v.2.1 multiple sequence alignment of putative VP/OT-type neuropeptides derived from precursors of representative species across the Bilateria. *B.flo_VP/OT*, *Branchiostoma floridae* VP/OT-type precursor [GI: 260828088]; *C.ele_VP/OT*, *Caenorhabditis elegans* VP/OT-type precursor [GI: 86564869]; *C.int_VP/OT*, *Ciona intestinalis* VP/OT-type precursor [GI: 209529695]; *C.tel_VP/OT*, *Capitella teleta* VP/OT-type precursor [JGI: 173251]; *D.pul_VP/OT*, *Daphnia pulex* VP/OT-type precursor [GI: 321460843]; *H.sap_OT*, *Homo sapiens* OT precursor [GI: 189410]; *H.sap_VP*, *Homo sapiens* VP precursor [GI: 340298]; *L.gig_VP/OT*, *Lottia gigantea* VP/OT-type precursor [GI: 676470613]; *S.kow_VP/OT*, *Saccoglossus kowalevskii* VP/OT-type precursor [GI: 187155721]; *S.pur_VP/OT*, *Strongylocentrotus purpuratus* VP/OT-type precursor [GI: 390337108]; *T.cas_VP/OT*, *Tribolium castaneum* VP/OT-type precursor [GI: 145651811]; *T.rub_OT*, *Takifugu rubripes* OT precursor [GI: 410922774]; *T.rub_VP*, *Takifugu rubripes* VP-type precursor [GI: 410922770].

3.1.1. VP/OT-type neuropeptide function in the vertebrates

VP and OT regulate a range of physiological processes in mammals and other vertebrates. In addition to the classic peripheral functions of VP in regulating vasoconstriction, osmoregulation (Henderson and Byron, 2007; McCormick and Bradshaw, 2006) and liver glycogenolysis (Eugenin et al., 1998; Michell et al., 1979) and OT in regulating reproductive processes (Gimpl and Fahrenholz, 2001), VP and OT have also been implicated with various central neuromodulatory functions.

VP and OT have been shown to have important roles in complex social cognition and behaviours including maternal-infant bonding and aggression (Broad et al., 2006; Caldwell et al., 2008; Donaldson and Young, 2008) in addition to pair bonding and attachment, which has been particularly well studied in vole species (Carter, 2014; Carter et al., 1995; Cho et al., 1999; Donaldson and Young, 2008; Insel and Young, 2001; Winslow et al., 1993; Young and Wang, 2004). These findings have stimulated a resurgence in interest in both VP and OT with respect to understanding human social behaviour. For example, VP and OT have recently been implicated in modulating social behaviour with respect to understanding autism, social anxiety disorder, borderline personality disorder and schizophrenia (McCall and Singer, 2012; Meyer-Lindenberg et al., 2011).

3.1.2. VP/OT-type neuropeptide function across the animal kingdom

Recent advances in comparative genomics/transcriptomics have provided new opportunities to discover neuropeptide signalling systems in a wider range of phyla and species across the animal kingdom. Despite the identification and characterisation of VP/OT-type neuropeptide signalling systems outside of the vertebrates, studies on the physiological roles that VP/OT-type peptides cover is

more limited. However, functional data has begun to provide important insights into the physiological roles of VP/OT-type peptides in these species and has provided insights into the evolutionarily ancient roles of the VP/OT-type neuropeptide signalling system in the common ancestor of the Bilateria.

3.1.2.1. VP/OT-type peptides in the deuterostomes

In the deuterostomian branch of the animal kingdom, recent genome/transcriptome sequencing has permitted the identification of VP/OT-type neuropeptide signalling systems in the urochordates, cephalochordates, hemichordates and the echinoderms. In the urochordates, VP/OT-type peptides have been identified in species including *Styela plicata* (Ukena et al., 2008), *C. intestinalis* (Kawada et al., 2008) and *Ciona savignyi* (Kawada et al., 2008). Functional studies in the urochordate *S. plicata* have revealed that *Styela* OT-related peptide (SOP) (CYISDCPNSRFWST-NH_2) causes contraction of *in vitro* inhalant and exhalant siphon preparations, suggesting that SOP has an osmoregulatory role in preventing the influx of a low concentration of seawater into the body (Ukena et al., 2008). VP/OT-type peptides have also been identified in the cephalochordate *B. floridae* (Elphick, 2010; Gwee et al., 2009) and the hemichordate *S. kowalevskii* (Elphick, 2010). However, to date, there is no functional data on the role of VP/OT-type peptides in these species. In the echinoderms, a VP/OT-type peptide ("echinotocin") has been identified in the sea urchin *Strongylocentrotus purpuratus* (Elphick and Rowe, 2009). Functional studies have revealed that echinotocin (CFISNCPKG-NH_2) causes contraction of *in vitro* tube foot and oesophagus preparations from the sea urchin species *Echinus esculentus*, suggesting that echinotocin may have an osmoregulatory role through the retraction of tube feet to reduce water influx in hypo-osmotic conditions (Elphick and Rowe, 2009). Taken together with the role of VP in vertebrates, data from the

urochordate *S. plicata* and the echinoderm *S. purpuratus* suggest that the VP/OT-type peptides may have an evolutionarily ancient role in osmoregulation (Elphick and Rowe, 2009).

3.1.2.2. VP/OT-type peptides in the protostomes

In the protostomian branch of the animal kingdom, recent genome/transcriptome sequencing has permitted the identification of VP/OT-type neuropeptide signalling systems across the lophotrochozoans including the mollusc *Lottia gigantea* (Veenstra, 2010) and the annelid *Capitella teleta* (Veenstra, 2011) and across the ecdysozoans including the arthropod *Tribolium castaneum* (Stafflinger et al., 2008) and the nematode *C. elegans* (Beets et al., 2012; Garrison et al., 2012). In comparison to the deuterostomian invertebrates, there have been more functional studies performed on the role of VP/OT-type peptides in the protostomian invertebrates. For example, functional studies in the mollusc *Lymnaea stagnalis* have revealed that Lys-conopressin (CFIRNCPKG-NH_2) induces the contraction of male copulatory organs (van Kesteren et al., 1995) and regulates glycogen metabolism (Geraerts, 1992; van Kesteren et al., 1995). Functional studies in the annelid *Whitmania pigra* have revealed that both annetocin (CFVRNCPTG-NH_2) and Lys-conopressin induce egg-laying-like behaviours and changes body weight as a measure of water balance, suggesting that both peptides have both reproductive and osmoregulatory functions respectively (Fujino et al., 1999; Oumi et al., 1996). Functional studies in the arthropod *T. castaneum* have revealed that an arginine VP-like peptide (AVPL) (CLITNCPRG-NH_2) has diuretic activity and acts indirectly to affect Malpighian tubule function, indicating that AVPL is involved in the regulation of water homeostasis in insects (Aikins et al., 2008). However, perhaps the most exciting breakthrough into the role of VP/OT-type peptides has come from recent

behavioural genetic studies in the nematode *C. elegans* in which central neuromodulatory functions of VP/OT-type peptides have been investigated. It has been shown that nematocin (CFLNSCPYRRY-NH_2) is involved in neuromodulation of male mating circuits and experience driven salt chemotaxis, a form of gustatory associative learning (Beets et al., 2012; Beets et al., 2013). It has therefore has been suggested that VP/OT-type peptides may in fact be ancient neuromodulators of neuronal circuits involved in associative learning (Beets et al., 2012; Beets et al., 2013). Taken together, functional data from the deuterostomes and the protostomes suggest that VP/OT-type peptides may have had an evolutionarily ancient role in both reproductive and osmoregulatory processes, whilst data from *C. elegans* has provided insights into potential roles in associative learning (Beets et al., 2012; Beets et al., 2013).

3.1.3. Investigating VP/OT-type peptides in *A. rubens*

The sequencing of the genome of the sea urchin *S. purpuratus* (Sodergren et al., 2006) led to the discovery of echinotocin, the first VP/OT-type peptide to be identified in an echinoderm species (Elphick and Rowe, 2009). However, despite echinotocin being shown to cause contraction of *in vitro* tube foot and oesophagus preparations in the sea urchin species *E. esculentus*, relatively little is known about the array of functions that echinotocin may cover in the sea urchin (Elphick and Rowe, 2009).

The VP/OT-type peptide ("asterotocin") in the common European starfish *Asterias rubens* (see **chapter 2.3.2.1**) is only the second VP/OT-type peptide to be identified in an echinoderm species and has provided further opportunities to investigate the function of VP/OT-type peptides in the echinoderms. The echinoderms provide a unique context for investigating the structure, function and

evolution of neuropeptide signalling systems; echinoderms possess a pentaradial morphological organisation along with the ability to autotomise, regenerate and rapidly and reversibly change the mechanical state of their body wall (Byrne, 2001; Patruno et al., 2001; Thorndyke et al., 2001; Wilkie, 2001; Wilkie, 2005). Moreover, as deuterostomian invertebrates, the echinoderms occupy an intermediate position in animal phylogeny, bridging model protostomian invertebrates (e.g. *Drosophila* and *C. elegans*) to the vertebrates. Therefore, investigating the function of asterotocin in *A. rubens* may provide important insights on neuropeptide diversity and its functional significance across the animal kingdom.

3.1.4. Aims and objectives

Herein, the analysis of neural transcriptome data from the starfish *A. rubens* has led to the discovery of a precursor protein that gives rise to a VP/OT-type peptide that has been termed "asterotocin" (see **chapter 2.3.2.1**). The discovery of the asterotocin precursor has provided an opportunity to:

(1) Confirm the sequence of the asterotocin precursor by cDNA cloning and sequencing.

(2) Determine the structure of the asterotocin peptide through the use of mass spectrometric methods.

(3) Investigate the expression of the asterotocin precursor in the starfish *A. rubens* using mRNA *in situ* hybridisation (mRNA ISH).

(4) Investigate the physiological functions of asterotocin in the starfish *A. rubens* using *in vivo* and *in vitro* pharmacological techniques.

3.2. Methods

3.2.1. Bioinformatic identification of the asterotocin precursor and a candidate asterotocin receptor in *A. rubens*

To search for a transcript encoding a VP/OT-type neuropeptide precursor in the starfish *A. rubens*, the sea urchin *S. purpuratus* VP/OT-type neuropeptide precursor ("echinotocin" precursor) protein sequence (Elphick and Rowe, 2009) was submitted as a query in a tBLASTn search of the *A. rubens* neural transcriptome using SequenceServer software (http://www.sequenceserver.com/) (Priyam et al., in prep). The protein sequence encoded by the top-hit (expect (e)-value) transcript was then analysed for the presence of an N-terminal signal peptide as predicted by SignalP 3.0 (Bendtsen et al., 2004), a VP/OT-type peptide flanked by dibasic/monobasic cleavage sites and a C-terminal neurophysin domain, which is characterised by the presence of fourteen cysteine (C) residues (de Bree, 2000; de Bree and Burbach, 1998). Subsequently, the protein sequence was submitted in a tBLASTn search of The National Centre for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine sequence similarity to known neuropeptide or peptide hormone precursors.

To search for a transcript encoding a candidate VP/OT-type receptor in the starfish *A. rubens*, the sea urchin *S. purpuratus* candidate echinotocin receptor protein sequence (Burke et al., 2006; Elphick and Rowe, 2009) was submitted as a query in a tBLASTn search of the *A. rubens* neural transcriptome using SequenceServer software (http://www.sequenceserver.com/) (Priyam et al., in prep). The protein encoded by the top-hit (e-value) transcript was then analysed for seven transmembrane domains characteristic of GPCRs using GPCRHMM (http://gpcrhmm.sbc.su.se). Subsequently, the protein sequence was submitted in a

tBLASTn search of The National Centre for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine sequence similarity to known neuropeptide or peptide hormone receptors.

3.2.2. Cloning and sequencing of the asterotocin precursor

Total RNA generated from the same specimen of *A. rubens* used for Illumina Hi-Seq (see **chapter 2.2.2**) was used to generate cDNA using the Quantitect Reverse Transcription Kit (QIAgen). Asterotocin precursor cDNA containing the entire coding region was amplified by PCR using Phusion High-Fidelity PCR Master Mix (NEB) and the oligos: 5'-ACACAGTACTACGATCAG-3' and 5'-GTCACAAGTGACCATATC-3'. The PCR product was gel-extracted and purified using the QIAquick Gel Extraction Kit (QIAgen) before being blunt-end cloned into a pBluescript SKII (+) vector (Agilent Technologies) cut with the EcoRV-HF (NEB) restriction endonuclease. Two clones were sequenced from both the T3 and T7 sequencing primer sites for sequence confirmation (Eurofins). A plasmid maxiprep was subsequently performed using the QIAgen Plasmid Maxi Kit (QIAgen).

3.2.3. Mass spectrometric confirmation of asterotocin

Radial nerve cords were dissected from five adult specimens of *A. rubens* using a method described previously (Chaet, 1964). Neuropeptides were then extracted in 1 ml of 80 % acetone on ice (Elphick et al., 1991a). Acetone was subsequently removed by evaporation using nitrogen, with the aqueous fraction centrifuged at 11,300 g in a MiniSpin® microcentrifuge (Eppendorf) for 10 minutes and the remaining supernatant stored at −80 °C. Prior to mass spectrometry (MS), the extract was thawed (with an aliquot diluted 10-fold with 0.1 % aqueous formic acid) and

filtered through a 0.22 µm Costar® Spin-X® centrifuge tube filter (Sigma-Aldrich) to remove particulates. The extract was analysed by nanoflow liquid chromatography (LC) with electrospray ionisation (ESI) quadrupole time-of-flight tandem MS (nano LC-ESI-MS/MS) using a nanoAcquity ultra performance LC (UPLC) system coupled to a Synapt® G2 High-Definition Mass Spectrometer™ (HDMS) (Waters Corporation) and MassLynx v.4.1 SCN 908 software (Waters Corporation). The mobile phases used for the chromatographic separation were 0.1 % aqueous formic acid (termed mobile phase A) and 0.1 % formic acid in acetonitrile (termed mobile phase B). An aliquot containing 15 µl of the *A. rubens* radial nerve cord extract was applied to a Symmetry C18® trapping column (180 µm x 20 mm, 5 µm particle size, 100 Å pore size) (Waters Corporation) using 99.9 % mobile phase A at a flow rate of 10 µl min^{-1} for 3 minutes, after which the fluidic flow path included the HSS T3 analytical capillary column (75 µm x 150 mm, 1.8 µm particle size, 100 Å pore size) (Waters Corporation). A linear gradient of 5–40 % mobile phase B over 105 minutes was utilised with a total run time of 120 minutes. The nanoflow ESI source conditions utilised a 3.5 kV capillary voltage, 25 V sample cone voltage and an 80 °C source temperature. The instrument was operated in resolution mode (~ 20,000 measured at full width and half height). A data-dependent acquisition was performed that would trigger an MS/MS scan on any singly charged peptide having a mass/charge (m/z) ratio of 960.3919, or a doubly charged peptide of m/z 480.6999, with a tolerance of 100 mDa allowed on the precursor m/z. MS/MS spectra, obtained from data-dependent acquisition, were processed using MassLynx v.4.1 SCN 908 software (Waters Corporation). Spectra were combined and processed using the MaxEnt 3 algorithm to generate singly charged, monoisotopic spectra for interpretation and manual validation.

3.2.4. Localisation of the asterotocin precursor in *A. rubens* using mRNA *in situ* hybridisation (mRNA ISH)

3.2.4.1. Probe synthesis

A total of 5 µg of maxipreped pBluescript SKII (+) vector containing the asterotocin precursor cDNA was linearised using either the restriction endonuclease SacI (NEB) (for sense probe) or KpnI (NEB) (for anti-sense probe) at approximately 3 hours at 37 °C. Linearised cDNA was cleaned using a phenol-chloroform (Sigma-Aldrich) and chloroform-isomylalcohol (Sigma-Aldrich) extraction. Asterotocin precursor cDNA was precipitated using 1/10 volume of 3 M sodium acetate and 2.5 volumes of 100 % ethanol at -80 °C for approximately 30 minutes before being vortexed for 30 minutes at 4 °C. The supernatant was then removed and the pellet was washed with 70 % ice-cold ethanol before air drying for approximately 1 hour and re-suspending in 12 µl autoclaved water. RNA probes were then synthesised using 2 µl digoxygenin nucleotide triphosphate (DIG-NTP) mix (Roche), 4 µl 5x transcription buffer (NEB), 2 µl 0.2 M dithiothreitol (DTT) (Promega) and 0.5 µl placental ribonuclease inhibitor (10 U/µl) (Promega) using 0.5 µl T7 polymerase (50 U/µl) (NEB) with 1 µg of SacI linearised cDNA (for sense probe) or 0.5 µl T3 polymerase (50 U/µl) (NEB) with 1 µg KpnI linearised cDNA (for anti-sense probe) in a 20 µl reaction for approximately 3 hours at 37 °C. Template DNA was then digested with RNase free DNase (NEB) for approximately 30 minutes at 37 °C. RNA probes were stored in 25% formaldehyde (FA) / 2x saline-sodium citrate (SSC) at -20 °C for long-term storage.

3.2.4.2. Tissue fixation

Small adult specimens of *A. rubens* were fixed in 4 % paraformaldehyde (PFA) in phosphate-buffered saline (PBS) overnight at 4 °C. Specimens were washed in PBS for approximately 10 minutes. *A. rubens* arms were dissected and placed in Morse's solution (10 % sodium citrate; 20 % formic acid in autoclaved water) for approximately 3 hours, whilst the central disk region was placed in Morse's solution for approximately 8 hours, with the Morse's solution changed approximately every 2-3 hours. *A. rubens* arms and central disk region were then washed in autoclaved water before being placed sequentially in 50 % ethanol, 70 % ethanol, 90 % ethanol and 100 % ethanol (x3) for approximately 30 minutes. *A. rubens* arms and central disk region were then placed in xylene solution for 5 minutes and then subsequently for 8 minutes. *A. rubens* arms and central disk region were then placed in melted filtered wax for approximately 1 hour (x3) to remove xylene before leaving overnight to solidify.

3.2.4.3. Tissue sectioning

A series of adjacent 14 µm sections of *A. rubens* arms and central disk (from at least three individual animals) were prepared on a RM 2145 microtome (Leica). Sections were collected on poly-L-lysine coated slides (VWR) covered with autoclaved water on a hot plate. Slides were left to dry for approximately 20 minutes before removing excess autoclaved water and leaving to completely dry overnight before proceeding with probe hybridisation and immunodetection.

3.2.4.4. Probe hybridisation and immunodetection

Slides were placed at 60 °C for approximately 1 hour to allow excess wax to melt before leaving to cool at room temperature for approximately 15 minutes. Slides were then placed in xylene solution for 7 minutes (x3) and then sequentially in 100 % ethanol (x2), 90 % ethanol, 70 % ethanol, 50 % ethanol, 30 % ethanol and then in PBS (x2) for 7 minutes respectively. Slides were then post-fixed in 4 % PFA/PBS for 20 minutes before washing with PBS/Tween 0.1 % for 5 minutes (x3) and subsequent treatment with 20 mg/ml proteinase K (QIAgen) (50 mM Tris-HCl [pH 7.5]; 6.25 mM EDTA in autoclaved water) at 37 °C for 12 minutes. Slides were then post-fixed in 4 % PFA/PBS for 5 minutes before washing with PBS/Tween 0.1 % for 5 minutes (x3) and then acetylated (1.325 % triethanolamine [pH 7-8]; 0.25 % acetic anhydride; 0.175 % HCl in autoclaved water) with stirring for 10 minutes. Slides were washed in PBS/0.1 % Tween-20 for 5 minutes (x2) and in 5x SSC for 5 minutes. Slides were then dried and placed in a humidified chamber containing 5x SSC and covered with hybridisation buffer (50 % formamide; 5x SSC; 500 µg/ml yeast; 50 µg/ml heparin; 0.1 % Tween-20 in autoclaved water) at room temperature for approximately 2 hours. Asterotocin precursor sense and asterotocin precursor anti-sense probes (800 ng/ml) were denatured in 100 µl of hybridisation buffer at 80 °C for 2 minutes and placed on ice before adding remaining hybridisation buffer to a total of 100 µl per slide. Slides were sealed with parafilm and placed in a humidified chamber containing 5x SSC at 65 °C overnight.

Parafilm was removed from slides by gentle shaking in 5x SSC heated to 65 °C and then washed in 0.2x SSC at 65 °C for 40 minutes (x2). Slides were then washed in 0.2x SSC at room temperature for 10 minutes and equilibrated in buffer B1 (10 mM Tris-HCl [pH 7.5]; 150 mM NaCl in autoclaved water) for 10 minutes. Slides were covered in buffer B1/5 % goat serum and placed in a humidified

chamber containing autoclaved water at room temperature for approximately 2 hours. Slides were then dried and covered in an alkaline phosphatase (AP)-conjugated anti-DIG antibody (1:3000; Roche) in buffer B1/ 2.5 % goat serum at 4 °C overnight. Slides were washed in buffer B1 for 5 minutes (x3) and then equilibrated in buffer B3 (100 mM Tris-HCl [pH 9.5]; 100 mM NaCl; 50 mM $MgCl_2$ in autoclaved water) for 10 minutes. Slides were then covered in buffer B3/0.1 % Tween-20 with nitro-blue tetrazolium chloride (NBT) (75 mg/ml in 70 % dimethylformamide) and 5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt (BCIP) substrate solution (50 mg/ml BCIP in autoclaved water) until desired staining was achieved. The staining reaction was stopped by rinsing slides in autoclaved water for 5 minutes (x3). Slides were then dried on a hot plate for approximately 45 minutes before rinsing for 10 seconds in 100 % ethanol (x2) and Histo-Clear (Natural Diagnostics) for 7 minutes (x2). Slides were finally mounted in Histo-Mount solution (Natural Diagnostics) for long-term storage.

Images were captured using a DMRA2 light microscope (Leica) with a MicroPublisher 5.0 Real-Time Viewing (RTV) digital colour camera (QImaging) and by utilising Volocity® v.5.3.1 image analysis software (PerkinElmer).

3.2.5. Analysis of the *in vivo* effects of asterotocin in *A. rubens*

Hamilton 75N 5 µl syringes (Sigma-Aldrich) were used to inject synthetic asterotocin (CLVQDCPEG-NH_2) (Peptide Protein Research Ltd.) at a range of concentrations into the perivisceral coelom of animals at two sites in the aboral body wall of the arms proximal to the junctions with the central disk region. Eversion of the cardiac stomach was observed in several animals. Subsequently, a total of thirty specimens of *A. rubens* were individually placed in a glass tank containing seawater. Animals were injected with 10 µl of 10 µmol l^{-1} synthetic asterotocin (a concentration selected based on preliminary observations) and video recorded for 15 minutes. A number of animals were first injected with 10 µl autoclaved water (control) and video recorded for 15 minutes to ensure cardiac stomach eversion was not due to injection alone. Care was taken to inject test agents into the perivisceral coelom and not into the cardiac stomach. Static images from video recordings of five animals in which the cardiac stomach was visible for the duration of the experiment were captured at 1-minute intervals from the time of injection. Then the two-dimensional area of everted cardiac stomach was measured from the images using ImageJ software (http://rsb.info.nih.gov/ij/) and normalised as a percentage of the area of the central disk of the animal.

3.2.6. Analysis of the *in vitro* effects of asterotocin on cardiac stomach preparations from *A. rubens*

Informed by the *in vivo* effect of synthetic asterotocin triggering cardiac stomach eversion, experiments were performed to investigate the *in vitro* effects of asterotocin on dissected cardiac stomach preparations. Cardiac stomachs were dissected from specimens of *A. rubens* and set up in a 20 ml organ bath maintained at 11 °C, as described previously (**Fig. 3.2**) (Elphick et al., 1995; Melarange et al., 1999). The organ bath contained artificial seawater supplemented with 30 mM potassium chloride (KCl) to maintain the cardiac stomach in a contracted state to allow for expected muscle relaxation to be recorded. Stock solutions of synthetic asterotocin (CLVQDCPEG-NH_2) (Peptide Protein Research Ltd.) were prepared in autoclaved water and then added to the organ bath to achieve final concentrations ranging from 30 pmol l^{-1} to 1 μmol l^{-1} upon the cardiac stomach maintaining a stabilised state. Cardiac stomach relaxation was recorded using an isotonic transducer (Harvard Apparatus) linked to a Goerz SE 120 chart recorder (Recorderlab). The preparation was subsequently washed with artificial seawater containing 30 mM KCl between different concentrations to allow the cardiac stomach preparation to return to a stabilised state.

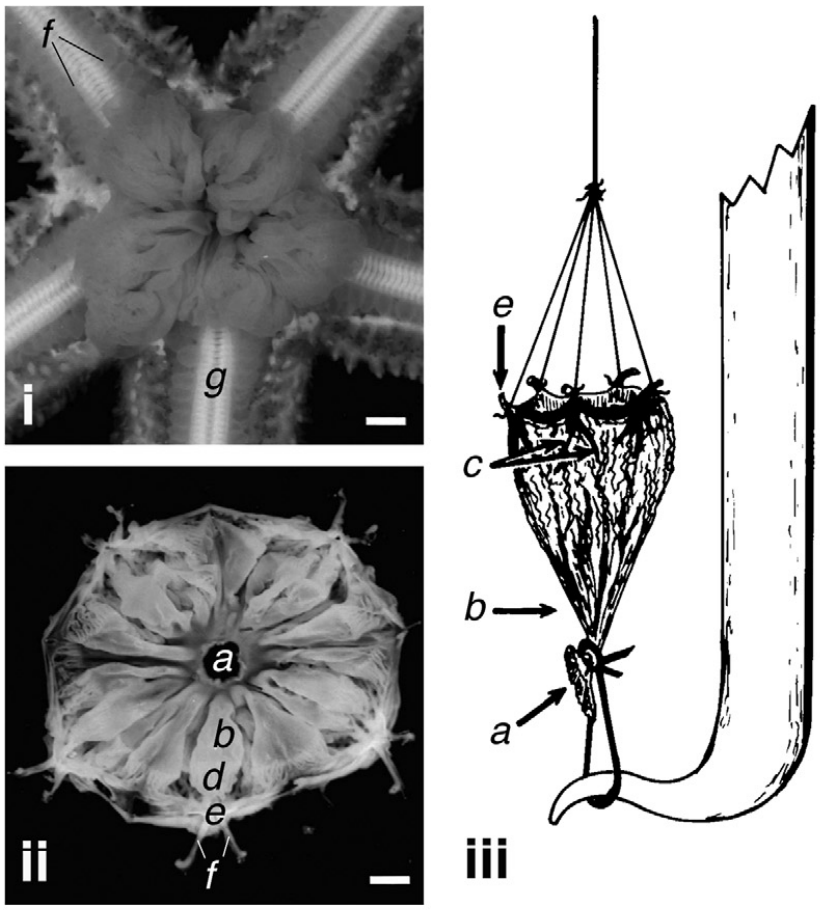


Fig. 3.2. Cardiac stomach assay. (i) Dissected specimen of *A. rubens* showing cardiac stomach *in situ*. **(ii)** An isolated *A. rubens* cardiac stomach *in vitro*. **(iii)** Illustration representing *in vitro* assay to measure cardiac stomach myoactivity. Cardiac stomach is tied to a force transducer and placed in an organ bath before addition of synthetic peptide to measure potential contraction and/or relaxation. *a*, oral opening; *b*, oesophagus; *c*, intrinsic retractor strands; *d*, aboral part of cardiac stomach; *e*, nodule; *f*, extrinsic retractor strand; *g*, ambulacrum. Scale bars, 0.5 cm. Figure from Elphick et al., 1995 and Elphick, 2014.

3.3. Results

3.3.1. Bioinformatic identification and cloning of the asterotocin precursor and identification of a candidate VP/OT-type receptor in *A. rubens*

In BLAST analysis of the *A. rubens* neural transcriptome data using the sea urchin *S. purpuratus* echinotocin precursor as a query, the top-hit (e-value: $1e^{-17}$) was contig 1119045 (2209 bp), which encodes a 147-residue protein termed the asterotocin precursor. The sequence of the asterotocin precursor has been confirmed by cDNA cloning and sequencing (**Fig. 3.3**). The precursor protein comprises a 23-residue N-terminal signal peptide (residues 1-23) followed by a 9-residue putative VP/OT-type peptide (CLVQDCPEGG) (residues 24-33), with the C-terminal glycine (G) residue predicted to be subject to amidation (asterotocin; CLVQDCPEG-NH_2). The asterotocin precursor also has a 78-residue C-terminal neurophysin domain comprising fourteen conserved cysteine (C) residues that are characteristic of VP/OT-type precursors (residues 44-121).

In BLAST analysis of the *A. rubens* neural transcriptome data using the sea urchin *S. purpuratus* VP/OT-type receptor as a query, the top-hit (e-value: $1e^{-114}$) was contig 1122053 (2710 bp), which encodes a 428-residue protein (**Fig. 3.4; Fig. A5L; Fig. S4**). The 428-residue protein contains seven transmembrane domains characteristic of GPCRs, with phylogenetic analyses revealing that the GPCR is a VP/OT-type receptor (**Fig. A2**).

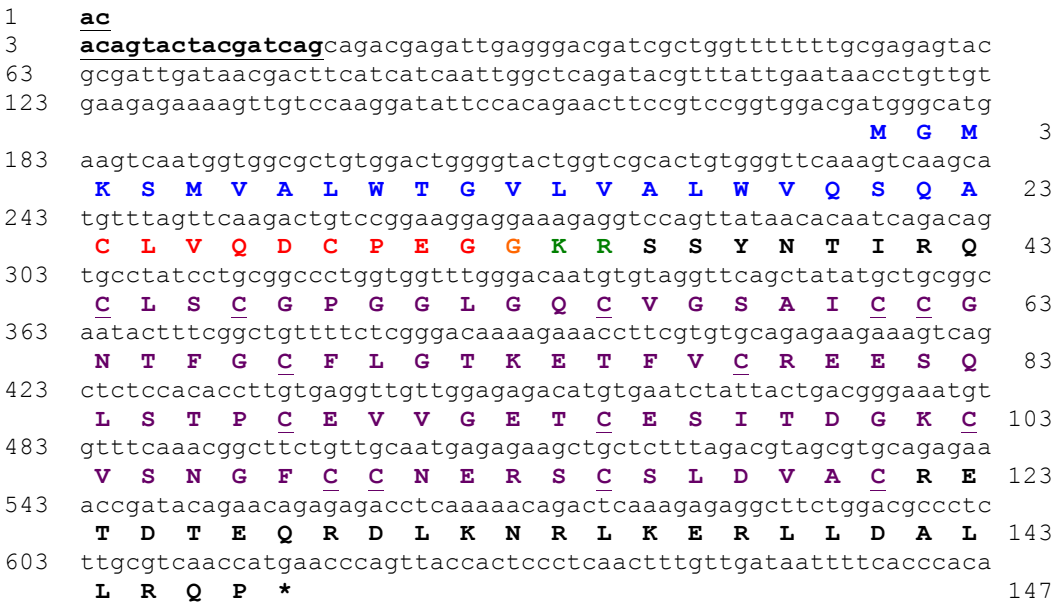


Fig. 3.3. ***A. rubens*** **asterotocin precursor cDNA sequence**. The nucleotide sequence (lowercase, 689 bases) encoding the asterotocin precursor protein (uppercase, 147 amino acid residues) is shown. Primers used for cloning the asterotocin precursor cDNA are represented in bold and underlined text. The predicted signal peptide is represented in blue, putative asterotocin peptide is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and a putative dibasic cleavage site (KR) is represented in green. The C-terminal region of the precursor comprises a neurophysin domain (with fourteen characteristic cysteine (C) residues underlined), which is represented in purple. The asterisk shows the position of the stop codon.

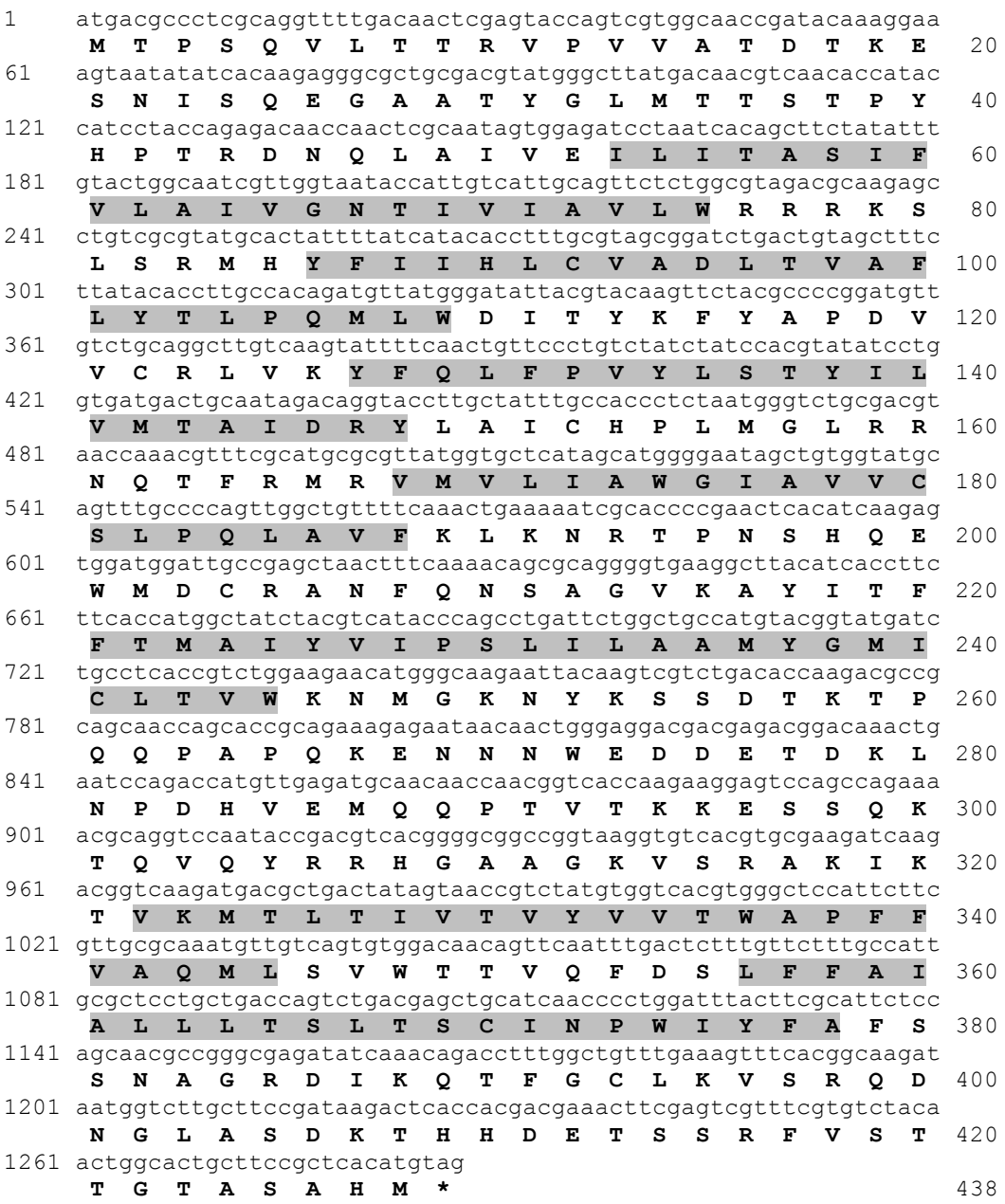
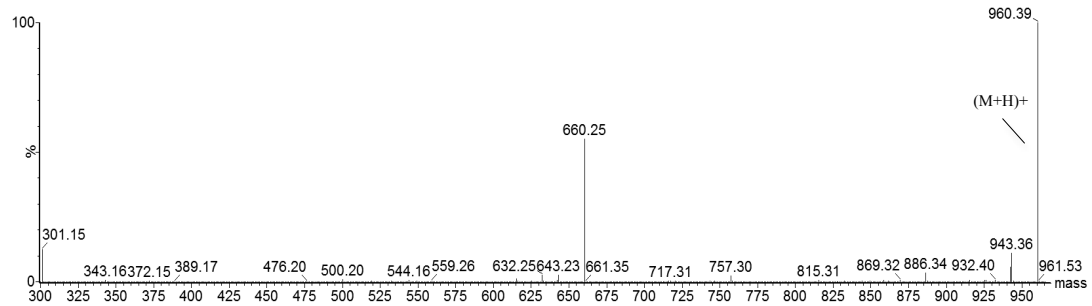


Fig. 3.4. ***A. rubens*** **VP/OT-type receptor.** The nucleotide sequence (lowercase, 1287 bases) encoding the receptor protein (uppercase, 438 amino acid residues) is shown. The asterisk shows the position of the stop codon. Predicted transmembrane domains are highlighted in grey.

3.3.2. Mass spectrometric confirmation of the structure of asterotocin

In order to perform both *in vivo* and *in vitro* functional studies, it is important to confirm the sequence and structure of the putative asterotocin peptide. Cloning and sequencing of the asterotocin precursor confirmed the presence of one copy of a VP/OT-type peptide ("asterotocin") with the sequence CLVQDCPEGG (**Fig. 3.3**). However, it is important to confirm whether the asterotocin peptide has a disulphide bridge between the cysteine (C) residues in positions 1 and 6 and whether the C-terminal glycine (G) residue is converted to an amide group. Nano LC-ESI-MS/MS using synthetic asterotocin as a standard was implemented, resulting in the peptide eluting at a retention time of 29.35 minutes with the singly charged species observed at a m/z of 960.39 (**Fig. 3.5A**). Subsequent analysis of the *A. rubens* radial nerve cord extract under identical conditions revealed that a single charged peptide with an m/z of 960.39 eluted at the same retention time to synthetic asterotocin (**Fig. 3.5B**). Both peptides were subjected to MS/MS during the experiment and the resulting deconvoluted, singly charged, monoisotopic spectra were compared, confirming the presence of asterotocin in the radial nerve cord extract (**Fig. 3.5A-B**).

(A)



(B)

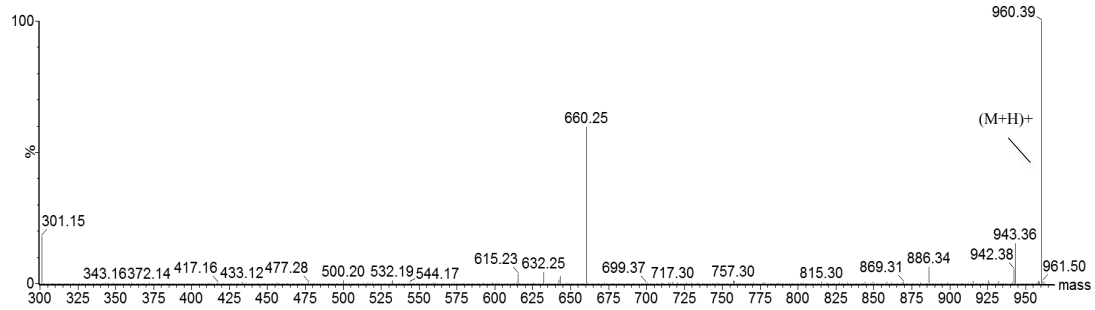


Fig. 3.5. Mass spectrometric confirmation that asterotocin is present in the radial nerve cord of *A. rubens*. (A) MS/MS spectrum of synthetic asterotocin peptide. **(B)** MS/MS spectrum of asterotocin in an acetone extract of radial nerve cords of *A. rubens*. The deconvoluted, monoisotopic, singly charged spectrum derived from MS/MS data is shown, with the asterotocin peptide ion annotated (960.39).

3.3.3. Expression profile of asterotocin mRNA in *A. rubens*

Serial sections were taken through the central disk and one arm of the starfish *A. rubens* to allow for visualisation of the expression of asterotocin precursor mRNA (or "asterotocin mRNA" for brevity) (**Fig. 3.6**). Expression of asterotocin mRNA was observed in the radial nerve cord, circumoral nerve ring, marginal nerve, tube feet, cardiac stomach, pyloric stomach, pyloric ducts, body wall and pedicellariae, which will be discussed on a one-by-one basis. Note that expression of asterotocin mRNA was not detected in control experiments utilising an asterotocin mRNA sense probe to detect non-specific binding (an example is provided in **Fig. 3.7**).

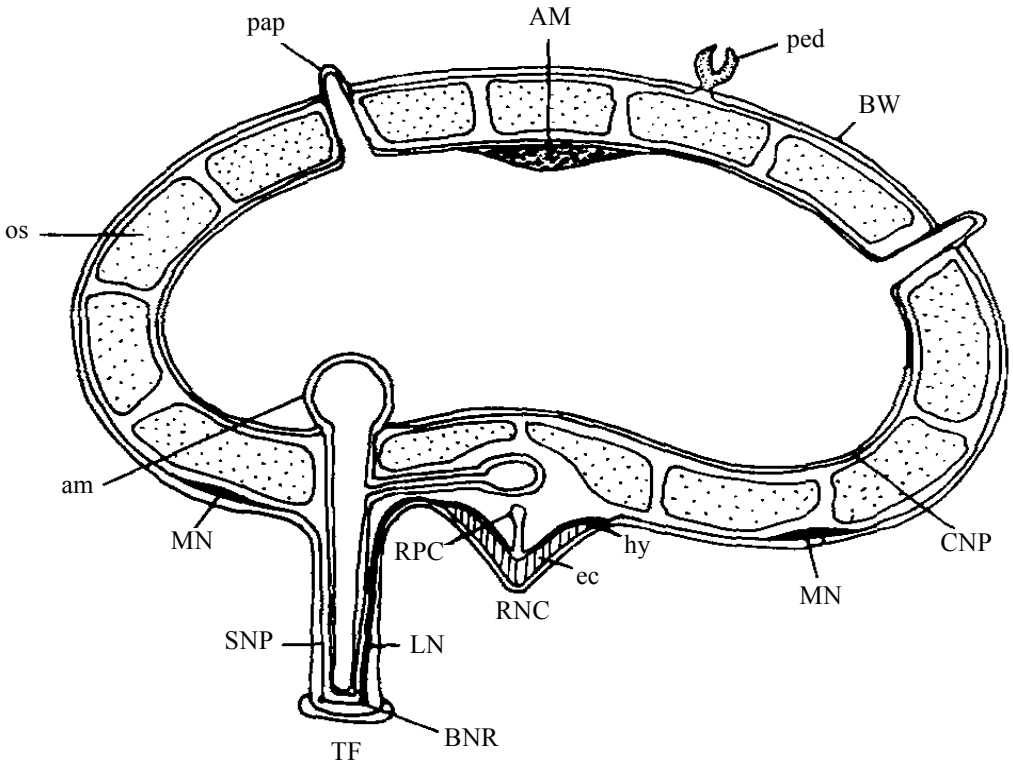


Fig. 3.6. Diagram of a transverse section of an arm of *A. rubens*. Ampulla (am); apical muscle (AM); basal nerve ring (BNR); body wall (BW); coelomic nerve plexus (CNP); ectoneural nerve plexus (ec); hyponeural nerve plexus (hy); longitudinal nerve (LN); marginal nerve (MN); ossicle (os); papulae (pap); pedicellariae (ped); radial nerve cord (RNC); radial perihaemal canal (RPC); sub-epithelial nerve plexus (SNP); tube feet (TF). Figure adapted from Moore and Thorndyke, 1993.

3.3.3.1. Radial nerve cord and circumoral nerve ring

The decentralised nervous system of the starfish *A. rubens* is comprised of five radial nerve cords, located in each arm between ambulacral ossicles, which connect to the circumoral nerve ring. The circumoral nerve ring is located in the central disk region surrounding the mouth and is attached to and innervates the peristomal epidermis (Moore and Thorndyke, 1993). The circumoral nerve ring and radial nerve cords comprise an ectoneural and a hyponeural region separated by a thin band of connective tissue (Cobb, 1970). The ectoneural nerve plexus is both sensory and motor in function, whilst the hyponeural nerve plexus is solely motor in function, innervating the tube feet and body wall muscles (Cobb, 1987; Moore and Thorndyke, 1993; von Hehn, 1970).

In transverse sections of the arm, expression of asterotocin mRNA was observed in the ectoneural region of the radial nerve cord but not in the hyponeural region (**Fig. 3.8.A-B**). The labelled cell bodies were located uniformally along the length of the radial nerve cord in the epithelium of the ectoneural nerve plexus (**Fig. 3.8.A-B**). Asterotocin mRNA expression was also present in cell bodies located in the lateral branches of the ectoneural region of the radial nerve cord (**Fig. 3.8D**). In sections through the central disk, expression of asterotocin mRNA was also observed in the ectoneural region of the circumoral nerve ring (**Fig. 3.8C**), consistent with the expression pattern observed in the ectoneural region of the radial nerve cord.

Fig. 3.7. Control experiment for asterotocin mRNA expression in the radial nerve cord in *A. rubens*. Transverse section through an arm of *A. rubens* displaying the radial nerve cord. Asterotocin mRNA expression was not detected in the radial nerve cord utilising an asterotocin mRNA sense probe to detect non-specific binding. *ec = ectoneural*; *hy = hyponeural*.

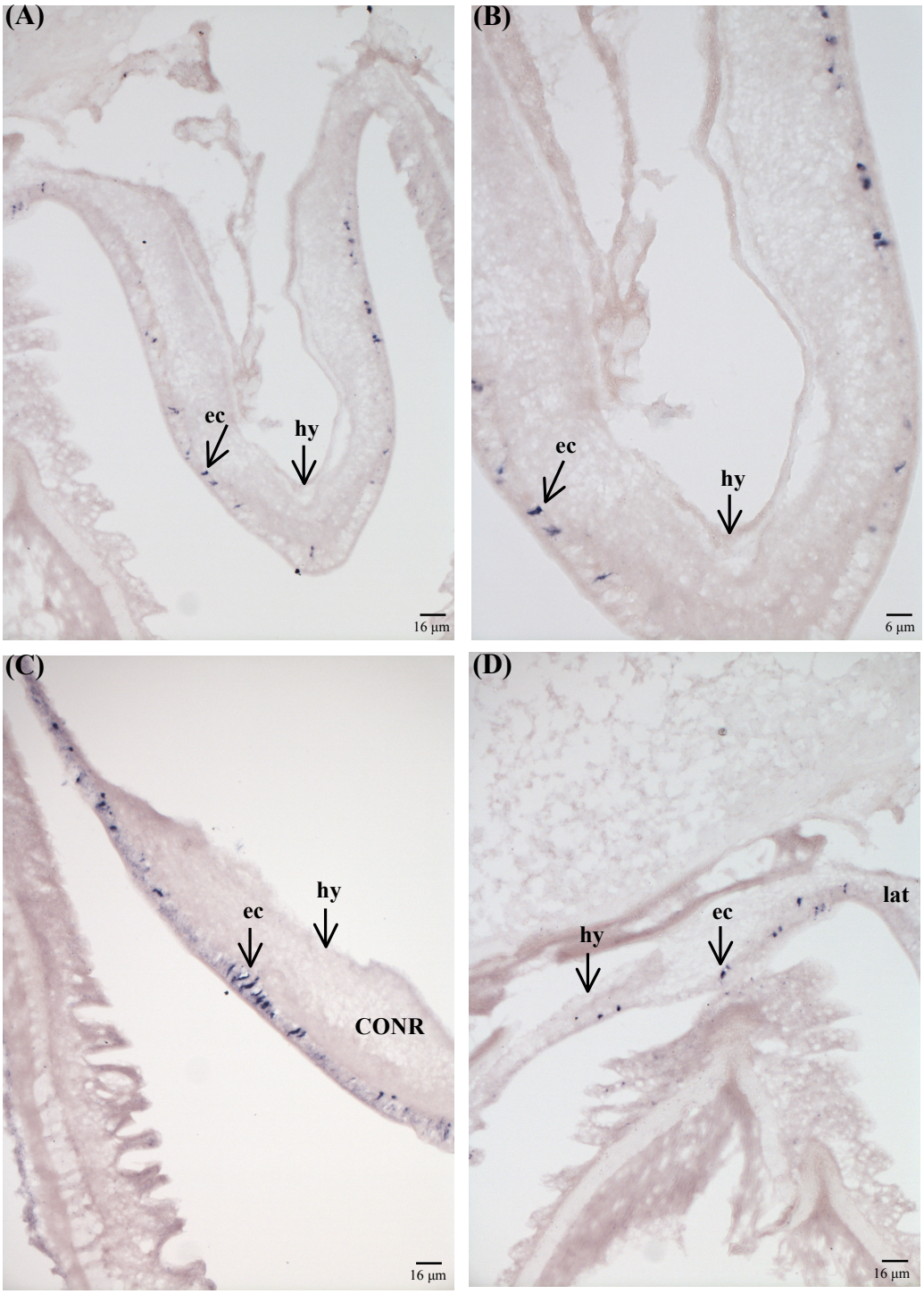

(A)
ec
hy
16 µm
(B)
ec
hy
6 µm
(C)
hy
ec
CONR
16 µm
(D)
lat
hy
ec
16 µm

Fig. 3.8. Asterotocin mRNA expression in the radial nerve cord, lateral branches of the radial nerve cord and circumoral nerve ring in *A. rubens*. (A-B) Transverse section through an arm of *A. rubens* displaying the radial nerve cord at two magnifications. Asterotocin mRNA expression was detected in the epithelium of the ectoneural nerve plexus of the radial nerve cord. **(C)** Transverse section through the central disk of *A. rubens* displaying the circumoral nerve ring. Asterotocin mRNA expression was detected in the ectoneural region of the circumoral nerve ring. **(D)** Transverse section through an arm of *A. rubens* displaying the lateral branches of the radial nerve cord. Asterotocin mRNA expression was detected in the ectoneural region of the lateral branches of the radial nerve cord. *CONR = circumoral nerve ring*; *ec = ectoneural*; *hy = hyponeural*; *lat = lateral branch of radial nerve cord*; *RNC = radial nerve cord*.

3.3.3.2. Tube feet and marginal nerve

Four rows of tube feet are located in the ambulacral groove, with a pair of rows of tube feet on either side of the radial nerve cord. The epidermis of the tube feet, with an underlying sub-epithelial nerve plexus, is continuous with the ectoneural region of the radial nerve cord (Moore and Thorndyke, 1993). The sub-epithelial nerve plexus is separated from longitudinally oriented retractor muscle fibres by a thin layer of connective tissue (Moore and Thorndyke, 1993). The sub-epithelial nerve plexus approaches the adhesive region of the tube foot (sucker) in a circular orientation and branches and thickens to form the basal nerve ring at the distal end of the tube foot stem (Newman et al., 1995a). The outer row of tube feet are innervated by a thickening of the body wall sub-epithelial nerve plexus called the marginal nerve, which is continuous with the extensive skin plexus (Moore and Thorndyke, 1993).

In transverse sections through the arm and central disk, expression of asterotocin mRNA was observed in cells located in the adhesive region of tube feet suckers (**Fig. 3.9A-B**). However, expression of asterotocin mRNA was not observed in the basal nerve ring or the sub-epithelial nerve plexus. Asterotocin mRNA expression was also detected in the marginal nerve, which innervates the outer row of tube feet (**Fig. 3.9C-D**).

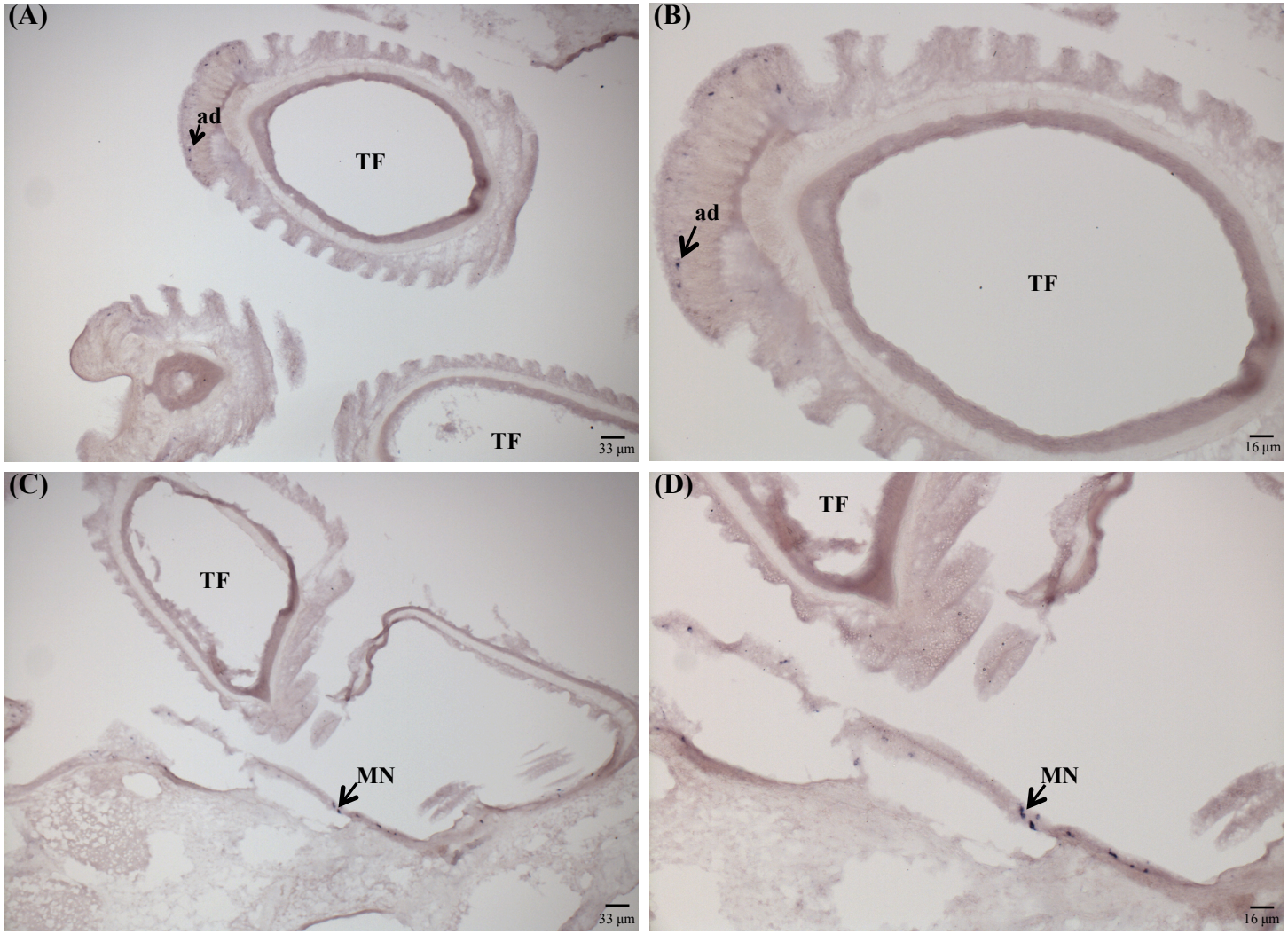

(A)
ad
TF
TF
33 µm
(B)
ad
TF
16 µm
(C)
TF
MN
33 µm
(D)
TF
MN
16 µm

Fig. 3.9. Asterotocin mRNA expression in the tube feet and marginal nerve in *A. rubens*. (A-B) Transverse section through an arm of *A. rubens* displaying tube feet at two magnifications. Asterotocin mRNA expression was detected in the nerve plexus in the adhesive region of the tube foot sucker. **(C-D)** Transverse section through a single arm of *A. rubens* displaying a marginal nerve supplying innervation to a tube foot at two magnifications. Asterotocin mRNA expression was detected in the marginal nerve. *ad = adhesive region of tube foot*; *MN marginal nerve*; *TF = tube foot*.

3.3.3.3. Digestive system

In the digestive system of the starfish *A. rubens*, the mouth is located at the centre of the disk on the oral side of the animal, surrounded by a muscular peristomal membrane. The mouth opens into a short oesophagus leading to the cardiac stomach (Ruppert et al., 2004). The cardiac stomach is an extensively folded structure that occupies a large proportion of the central disk and is connected aborally to a smaller pyloric stomach. A short duct, termed the pyloric duct, connects the pyloric stomach to a pair of digestive glands (pyloric caecae) in each arm. The aboral side of the pyloric stomach is connected to the anus by short tubular intestine (rectum), which has diverticulae known as rectal caecae (Ruppert et al., 2004).

The gut is comprised of a single layer of ciliated peritoneal epithelial cells, with two thin layers of longitudinal and circular muscle (comprising the visceral muscle layer) associated innervation beneath (**Fig. 3.10**) (Moore and Thorndyke, 1993). The visceral muscle layer is separated from the basiepithelial nerve plexus and adjacent epithelial cells lining the lumen of the gut by a continuous layer of connective tissue (Moore and Thorndyke, 1993). The majority of cells lining the lumen of the gut are ciliated, allowing the movement of partially digested material from the cardiac stomach through to the pyloric caeca. However, other cell types lining the lumen of the gut include mucosal goblet cells and granular secretory cells (Moore and Thorndyke, 1993).

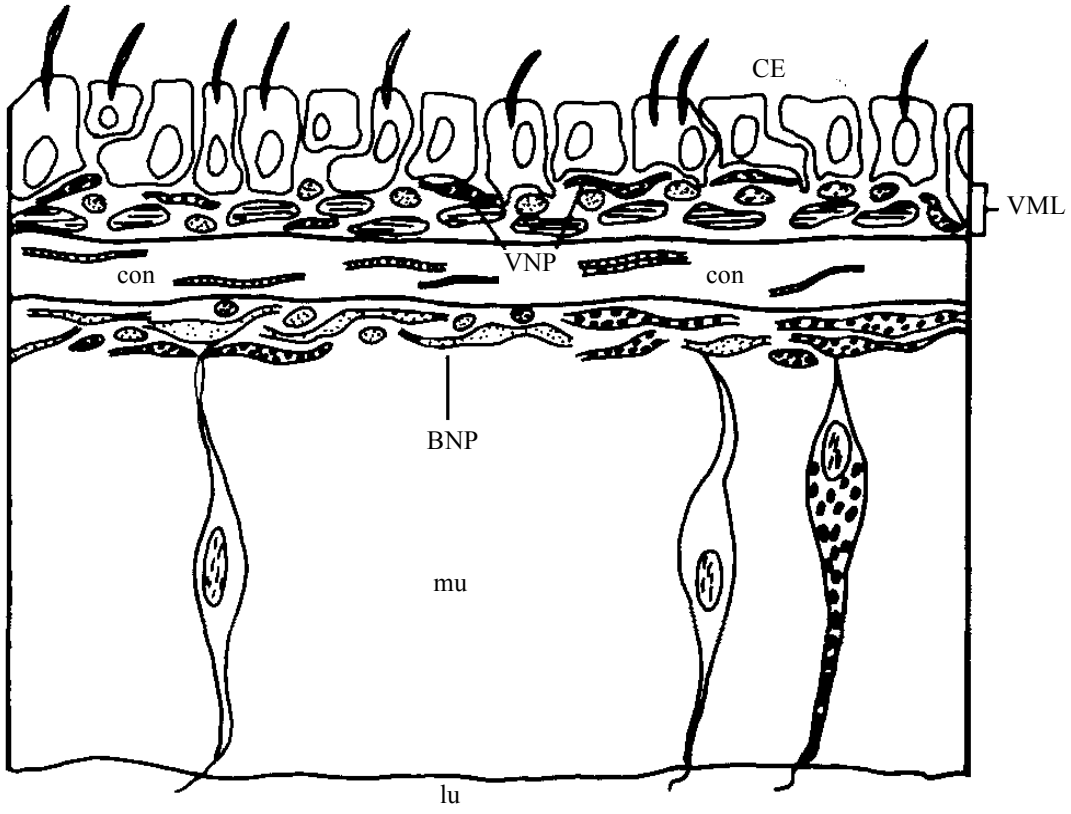


Fig. 3.10. Diagram of the anatomy of the gut of *A. rubens*. Basiepithelial nerve plexus (BNP); coelomic epithelium (CE); connective tissue (con); lumen (lu); mucosa (mu); visceral muscle layer (VML); visceral nerve plexus (VNP). Figure adapted from Moore and Thorndyke, 1993.

In sections through the central disk, expression of asterotocin mRNA was observed in the cardiac stomach (**Fig. 3.11**) with sparser expression detected in the pyloric stomach and very sparse expression detected in the pyloric ducts (**Fig. 3.12A-B** and **Fig. 3.12D**). However, expression of asterotocin mRNA was not observed in the peristomial epithelium, oesophagus, pyloric caecae (**Fig. 3.12C**), rectum, rectal caecae (**Fig. 3.12A**) or retractor strands. Labelled cell bodies appear to be more densely concentrated in the aboral region as opposed to the oral region of the cardiac stomach, which is densely packed with granular secretory cells (**Fig. 3.11**). In both the aboral and oral regions of the cardiac stomach, however, labelled cell bodies are largely concentrated in the basiepithelial nerve plexus (**Fig 3.11**). Labelled cell bodies are also located in the mucosal layer, with very sparse expression in the visceral muscle layer (**Fig. 3.11**). In the pyloric stomach, labelled cell bodies are sparsely located in the basiepithelial nerve plexus but are absent from the mucosal and visceral muscle layers (**Fig. 3.12B**). In the pyloric ducts, cell bodies are sparsely located in the basiepithelial nerve plexus and the mucosal layer but are absent from the visceral muscle layer (**Fig. 3.12D**).

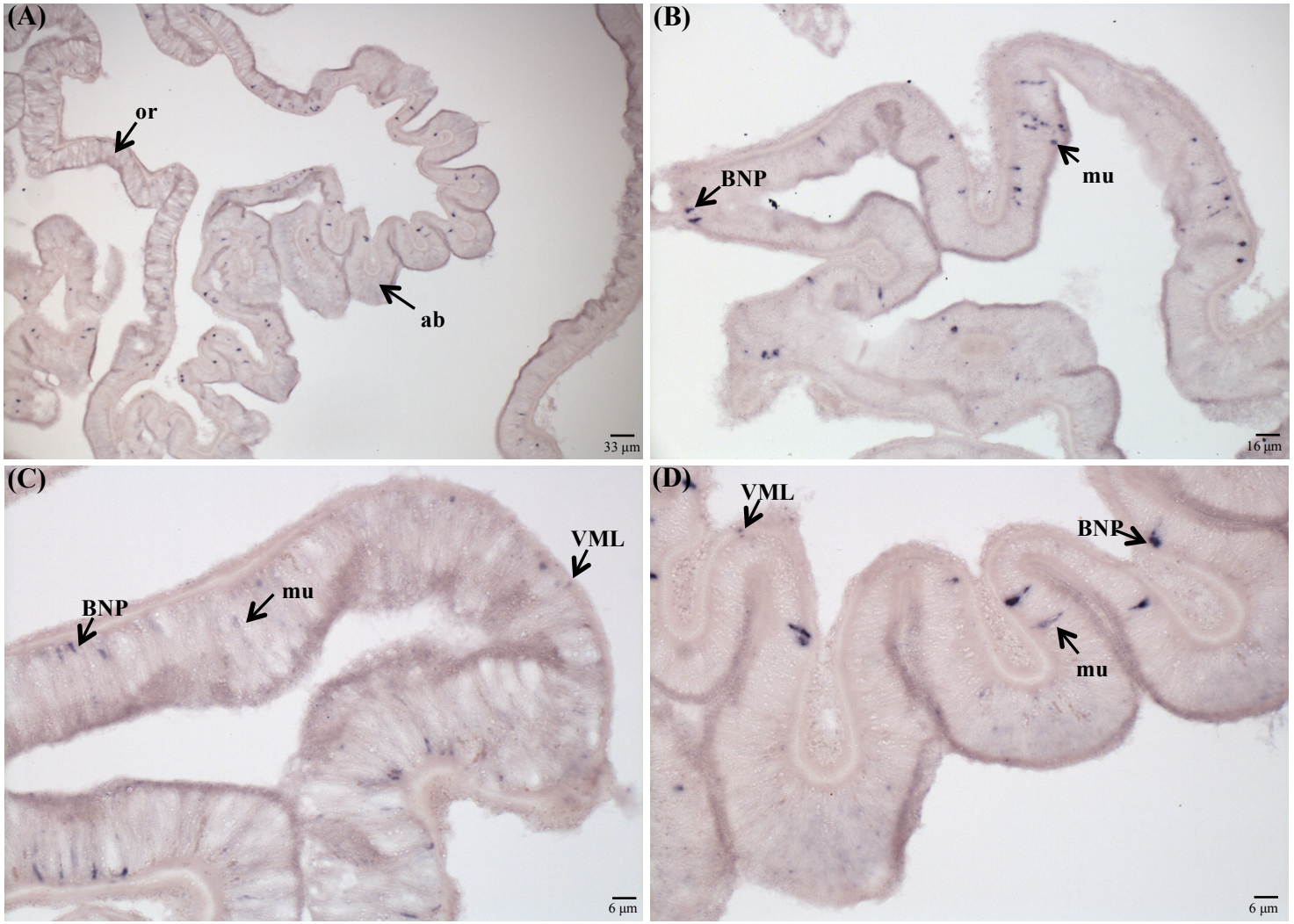

(A)
or
ab
33 µm
(B)
BNP
mu
16 µm
(C)
VML
BNP
mu
6 µm
(D)
VML
BNP
mu
6 µm

Fig. 3.11. Asterotocin mRNA expression in the cardiac stomach of *A. rubens*. (A) Transverse section through the central disk of *A. rubens* displaying the oral and aboral regions of the cardiac stomach. **(B)** Transverse section through the central disk of *A. rubens* displaying the aboral region of the cardiac stomach. Asterotocin mRNA expression was detected in the basiepithelial nerve plexus, with sparse expression in the mucosal layer. **(C)** Transverse section through the central disk of *A. rubens* displaying an oral region of the cardiac stomach. Asterotocin mRNA expression was detected in the basiepithelial nerve plexus, with sparse expression in the mucosal layer and the visceral muscle layer. **(D)** Transverse section through the central disk of *A. rubens* displaying the aboral region of the cardiac stomach. Asterotocin mRNA expression was detected in the basiepithelial nerve plexus, with sparse expression in the mucosal layer and visceral muscle layer. *ab = aboral cardiac stomach; BNP = basiepithelial nerve plexus*; *mu = mucosal layer*; *or = oral cardiac stomach*; *VML = visceral muscle layer*.

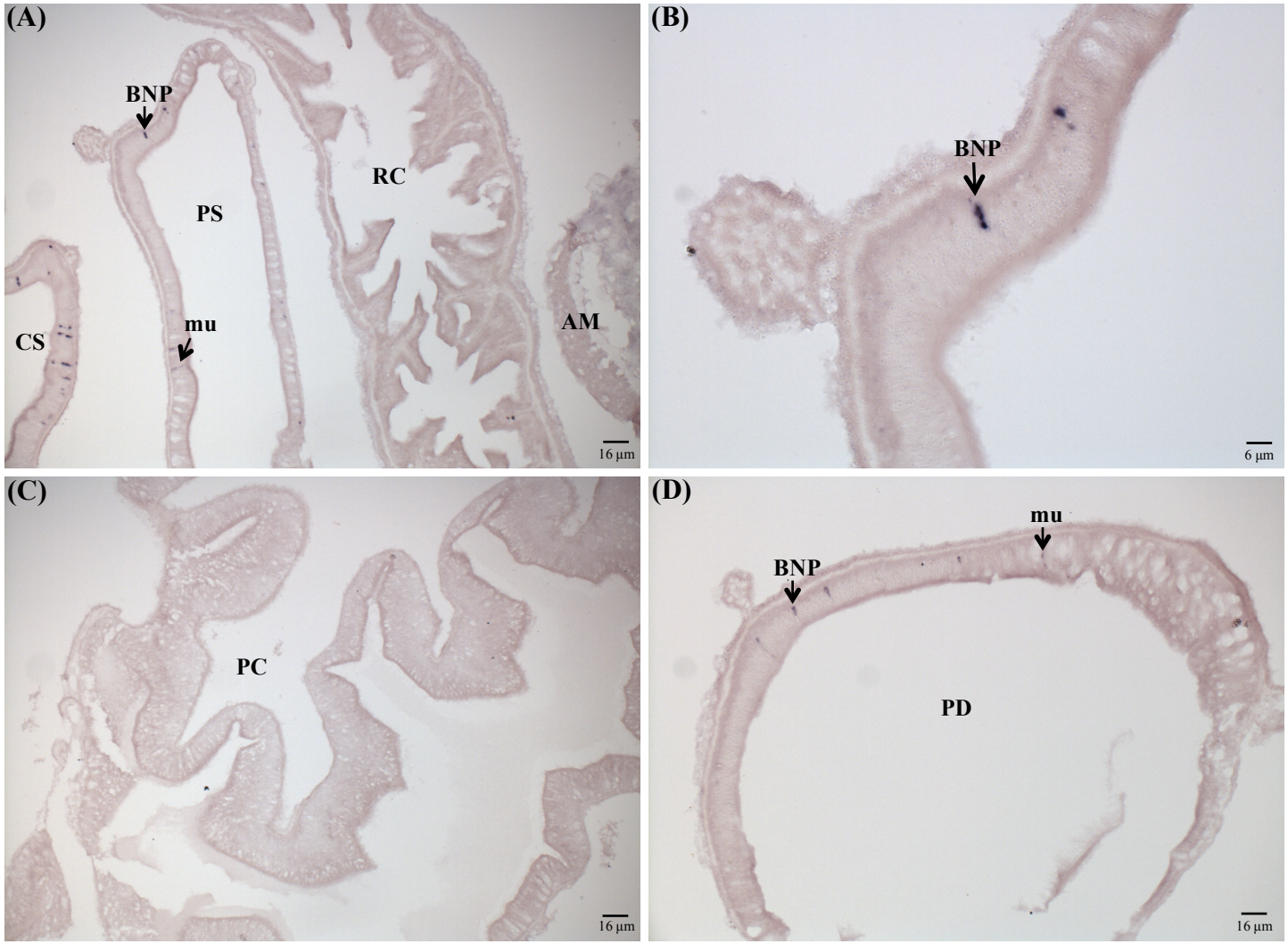

(A)
BNP
RC
PS
CS
mu
AM
16 µm
(B)
BNP
6 µm
(C)
PC
16 µm
(D)
mu
BNP
PD
16 µm

Fig. 3.12. Asterotocin mRNA expression in regions of the digestive system in *A. rubens*. **(A)** Transverse section through the central disk of *A. rubens* displaying the pyloric stomach and rectal caecae. Asterotocin mRNA expression was detected in the basiepithelial nerve plexus and mucosal layer of the pyloric stomach. **(B)** Transverse section through the central disk of *A. rubens* displaying the pyloric stomach. Asterotocin mRNA expression was detected in the basiepithelial nerve plexus and mucosal layer of the pyloric stomach. **(C)** Transverse section through the central disk of *A. rubens* displaying pyloric caecae. Asterotocin mRNA expression was not detected in the pyloric caecae. **(D)** Transverse section through the central disk of *A. rubens* displaying a pyloric duct. Asterotocin mRNA expression was detected in the basiepithelial nerve plexus and mucosal layer of the pyloric duct. *AM = apical muscle*; *BNP = basiepithelial nerve plexus*; *CS = cardiac stomach*; *mu = mucosal* layer; PC = *pyloric caecae*; *PD = pyloric duct*; *PS = pyloric stomach*; *RC = rectal caecae*.

3.3.3.4. Body wall and pedicellariae

The starfish body wall comprises an epidermal layer, a dermal layer in which calcite ossicles are embedded, two layers of longitudinal and circular muscle forming the visceral muscle layer and an inner coelomic epithelium (O'Neill, 1989). The dermal layer is packed with fibres comprised of collagen fibrils forming an orthogonal array (Motokawa, 2011; O'Neill, 1989). In between these collagen fibrils, juxtaligamental cells controlling the mechanical properties of connective tissue are found (Motokawa, 2011; Wilkie, 2005). The expression of asterotocin mRNA was observed in cell bodies in the epidermal layer of the body wall (**Fig. 3.13A**). However, asterotocin mRNA was not observed in the dermal layer or in the inner coelomic epithelium (**Fig. 3.13A**).

The pedicellariae are small jaw-like structures present on the aboral surface that are primarily used to prevent colonisation of microorganisms (Roberts and Campbell, 1988). The pedicellariae consist of a short stalk surpassed by small ossicles, which together form the characteristic jaw-like structure (Moore and Thorndyke, 1993). The expression of asterotocin mRNA was observed in cell bodies in the epidermal layer of the pedicellariae, which lines both inside and outside of the jaw (**Fig. 3.13B**).

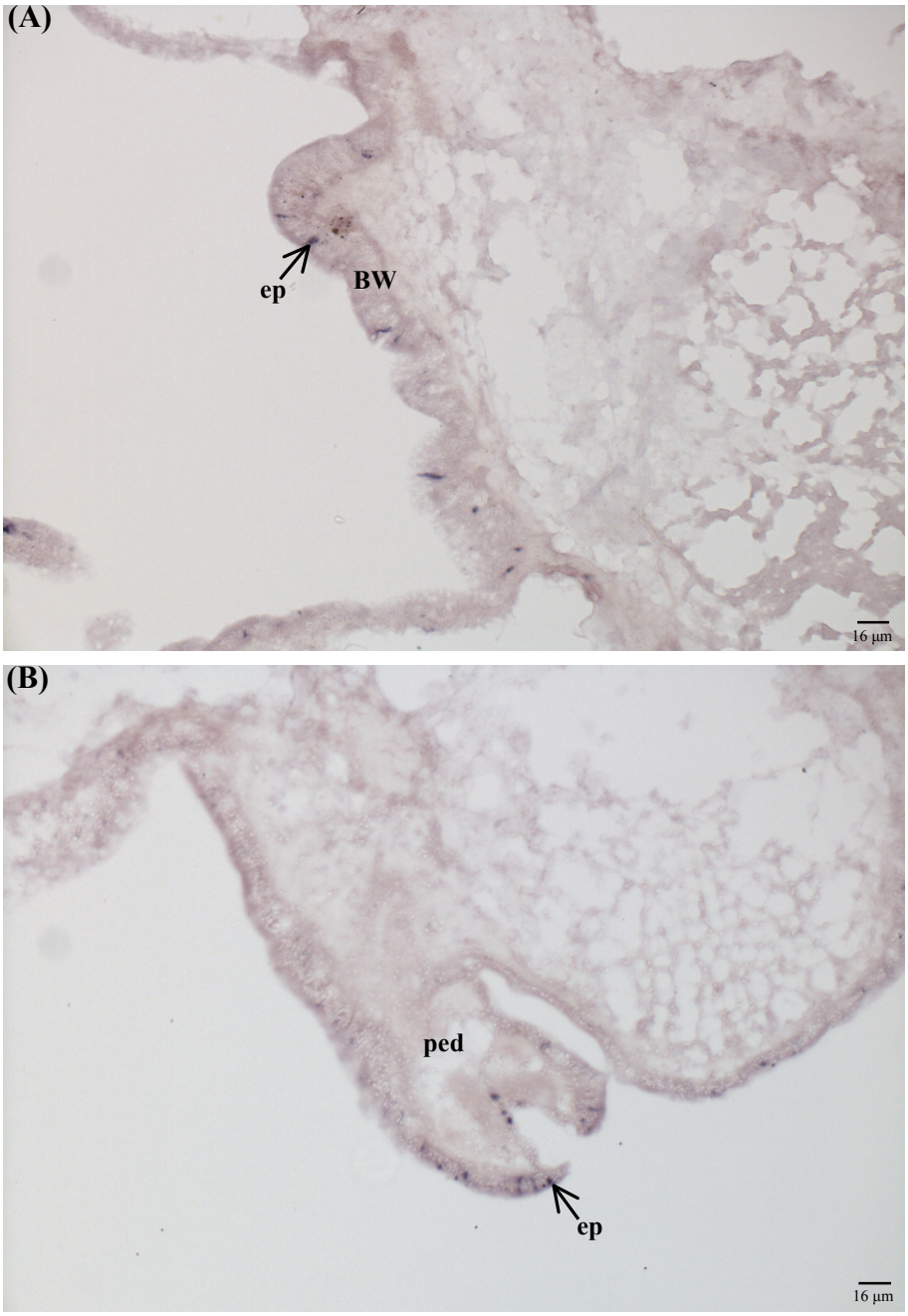
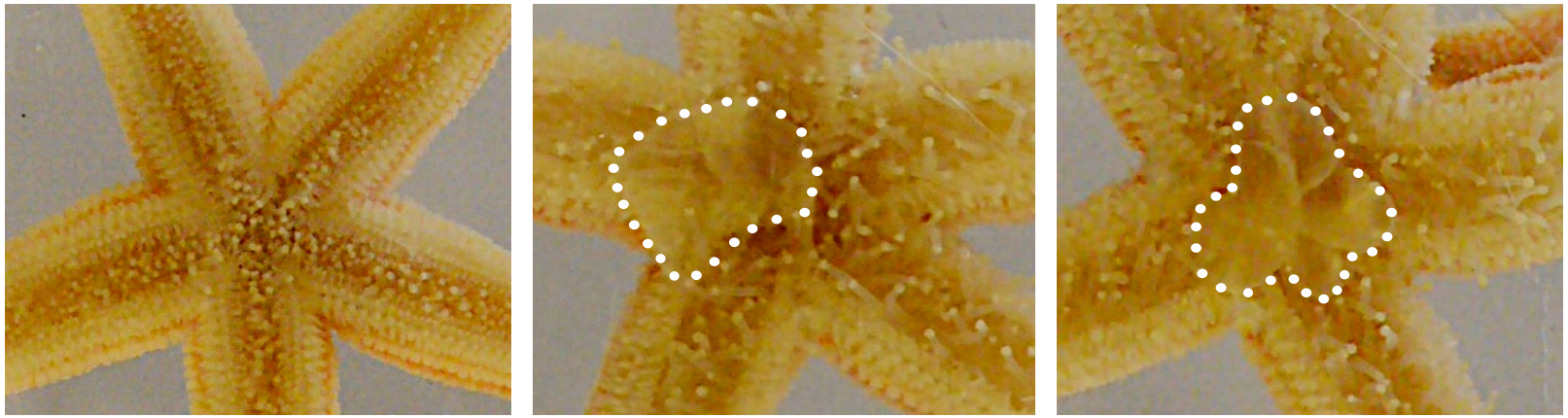


Fig. 3.13. Asterotocin mRNA expression in the body wall and pedicellariae in *A. rubens*. (A) Transverse section through the central disk of *A. rubens* displaying the body wall. Asterotocin mRNA expression was detected in cell bodies in the epidermal layer of the body wall. **(B)** Transverse section through the central disk of *A. rubens* displaying a single pedicellariae. Asterotocin mRNA expression was detected in cell bodies in the epidermal layer of the pedicellariae. *BW = body wall*; *ep = epidermal layer*; *ped = pedicellariae*.

3.3.4. Asterotocin causes cardiac stomach eversion *in vivo*

In order to determine whether synthetic asterotocin (CLVQDCPEG-NH_2) has an observable effect on any physiological processes *in vivo*, asterotocin was injected into the perivisceral coelom of the central disk region of the starfish *A. rubens*. Injections of the SALMFamides S1 (GFNSALMF-NH_2) and S2 (SGPYSFNSGLTF-NH_2) have previously been shown to trigger cardiac stomach eversion in the starfish *A. rubens* (Elphick et al., 1995; Melarange et al., 1999). Interestingly, in the *in vivo* assay, injection of asterotocin (10 µl of 10 µmol l^{-1}) into the perivisceral coelom also triggered eversion of the cardiac stomach. Informed by this initial observation, injection experiments were performed on a total of thirty specimens in order to determine whether the effect of asterotocin (10 µl of 10 µmol l^{-1}) on cardiac stomach eversion was repeatable. Asterotocin triggered cardiac stomach eversion in 17/30 experiments with variability in the rate and extent of eversion (**Fig. 3.14**). **Fig. 3.14B** shows data from five experiments in which the cardiac stomach could be monitored for the full duration of the experiment, with the mean area of cardiac stomach everted at 1-minute intervals during a 15 minute recording period following peptide injection at time 0 (T_0) expressed as a percentage of the area everted in relation to the central disk area. Importantly, in control experiments in which starfish were injected with autoclaved water, no eversion of the cardiac stomach was observed. It should also be noted that in a number of specimens, postural changes including the thinning of arms, curling of arms and tube foot extension could be observed (**Fig. 3.15**).

(A)


0 min
7.5 min
15 min

(B)

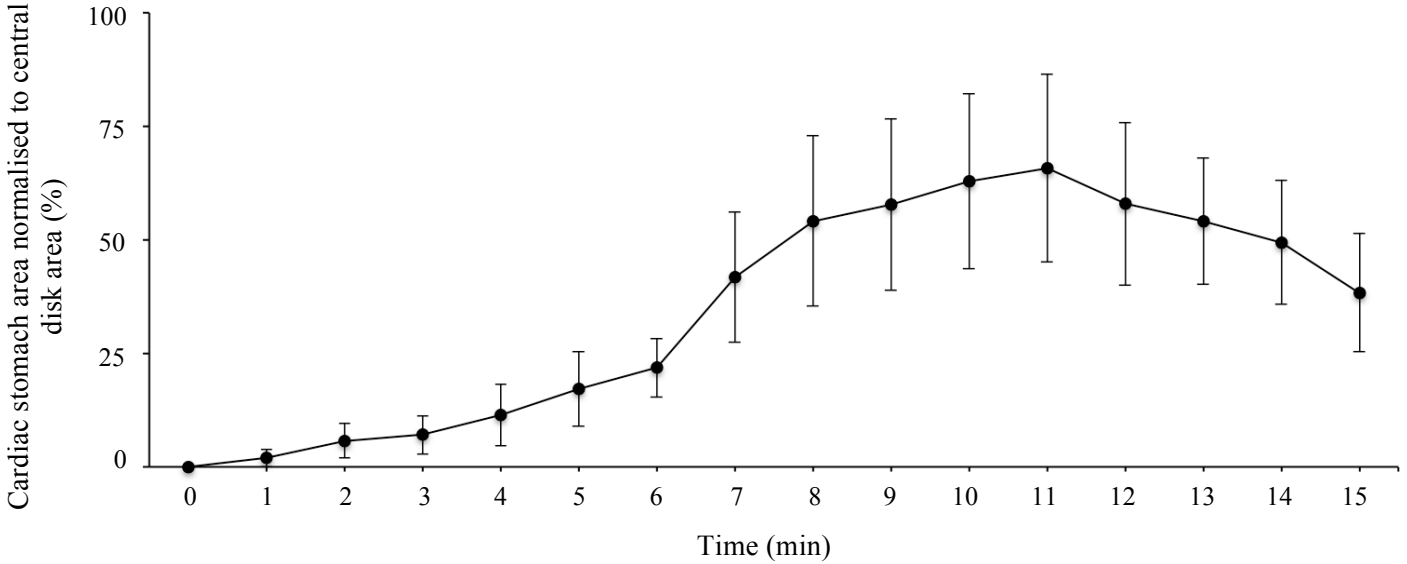

Cardiac stomach area normalised to central disk area (%)
100
75
50
25
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Time (min)

Fig. 3.14. Asterotocin triggers cardiac stomach eversion in *A. rubens*. **(A)** Photographs from a representative experiment showing that injection of synthetic asterotocin (10 µl of 10 µmol l^{-1}) causes cardiac stomach eversion. At time 0 (T_0), no cardiac stomach eversion can be seen. At 7.5 minutes and 15 minutes after injection of asterotocin, cardiac stomach eversion is observed (outline of cardiac stomach marked by white dots). **(B)** Graph showing that injection of synthetic asterotocin (10 µl of 10 µmol l^{-1}) causes cardiac stomach eversion. The area of cardiac stomach everted (in 2D) at each time point (0-15 minutes) is normalised to the area of central disk, with means (± s.e.m.) from five experiments shown.

Fig. 3.15. Representative example of postural changes in the starfish *A. rubens* following injection of synthetic asterotocin. Following injection of synthetic asterotocin (10 µl of 10 µmol l^{-1}), in addition to cardiac stomach eversion (not shown), postural changes including the thinning of arms, curling of arms and tube foot extension was observed.

3.3.5. Asterotocin causes cardiac stomach relaxation *in vitro*

In order to determine whether the effect of synthetic asterotocin on cardiac stomach eversion *in vivo* (**Fig. 3.14**) correlates with an effect on cardiac stomach relaxation *in vitro*, an *in vitro* assay previously used to show that the SALMFamides S1 and S2 cause cardiac stomach relaxation was implemented (**Fig. 3.2**) (Elphick and Melarange, 1998; Elphick et al., 1995; Melarange et al., 1999). In the *in vitro* assay, the addition of asterotocin to cardiac stomach preparations revealed that it caused dose-dependent relaxation at concentrations ranging from 30 pmol l^{-1} to 1 μmol l^{-1} (**Fig. 3.16**); this is consistent with the effects of asterotocin on cardiac stomach eversion *in vivo*. A total of sixteen adult specimens were analysed, with the maximal efficacy of asterotocin observed at 100 nmol l^{-1} (**Fig. 3.16**).

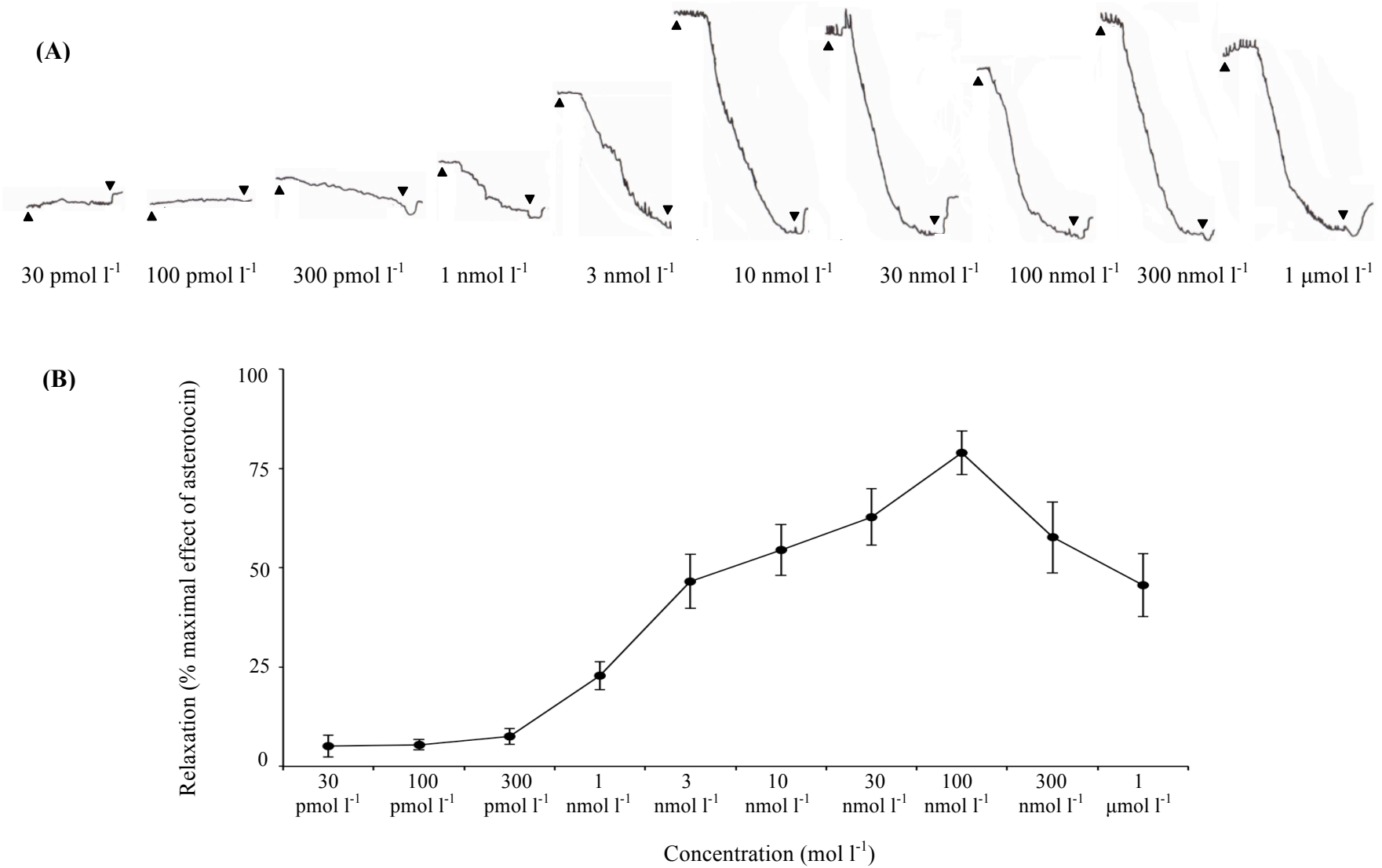

(A)
30 pmol l-1
100 pmol l-1
300 pmol l-1
1 nmol l-1
3 nmol l-1
10 nmol l-1
30 nmol l-1
100 nmol l-1
300 nmol l-1
1 µmol l-1
(B)
Relaxation (% maximal effect of asterotocin)
100
75
50
25
0
30 pmol l-1
100 pmol l-1
300 pmol l-1
1 nmol l-1
3 nmol l-1
10 nmol l-1
30 nmol l-1
100 nmol l-1
300 nmol l-1
1 µmol l-1
Concentration (mol l-1)

Fig. 3.16. Asterotocin is a potent stimulator of cardiac stomach relaxation in *A. rubens*. **(A)** Representative recordings from a single cardiac stomach preparation showing the dose-dependent effect of synthetic asterotocin (30 pmol l^{-1} to 1 μmol l^{-1}). The application of asterotocin (upward-pointing arrowheads) causes cardiac stomach relaxation in a typical time-frame of 1-2 minutes, whilst the effect of asterotocin is reversed by washing (downward-pointing arrowheads). **(B)** Graph showing the dose-dependent effects of synthetic asterotocin (30 pmol l^{-1} to 1 μmol l^{-1}). The effects of asterotocin are normalised to the maximal effect observed with asterotocin in each experiment, with mean values (± s.e.m.) from sixteen experiments shown.

3.4. Discussion

3.4.1. Discovery of asterotocin, a VP/OT-type peptide in *A. rubens*

The asterotocin precursor in the starfish *A. rubens* is only the second VP/OT-type precursor to be discovered in an echinoderm species after the discovery of the echinotocin precursor in the sea urchin *S. purpuratus* (Elphick and Rowe, 2009). The asterotocin precursor contains an N-terminal signal peptide, one copy of the sequence CLVQDCPEGG flanked by a dibasic cleavage site (KR) and a C-terminal neurophysin domain (**Fig. 3.3**). Importantly, mass spectrometry has confirmed that the C-terminal glycine (G) residue of asterotocin is converted to an amide group (CLVQDCPEG-NH_2) (**Fig. 3.5**). In addition to the presence of a C-terminal amide group, comparison of asterotocin with other VP/OT-type peptides in phyla across the animal kingdom reveals interesting similarities and differences. Asterotocin, as with all known members of the VP/OT neuropeptide family, comprises two conserved cysteine (C) residues in positions 1 and 6 of the mature neuropeptide that has been shown to form a disulphide bridge and confer a cyclic conformation that is important for biological activity (Hruby et al., 1990; Sawyer, 1977). Importantly, the presence of this disulphide bridge has also been confirmed by mass spectrometry (**Fig. 3.5**). Asterotocin has a leucine (L) and a valine (V) residue in positions 2 and 3 of the mature neuropeptide, respectively, which is consistent with the occurrence of hydrophobic residues in these positions in other members of the VP/OT neuropeptide family. Moreover, asterotocin has an aspartic acid (D) residue in position 5 of the mature neuropeptide similar to two VP/OT-type peptides identified in the urochordates *C. intestinalis* and *S. plicata* (Kawada et al., 2008; Ukena et al., 2008) and in the nematode *C. elegans* (Beets et al., 2012; Garrison et al., 2012). Asterotocin also has a proline (P) residue in position 7 along with a C-terminal

glycine (G) residue in position 9, which undergoes C-terminal amidation as with all other known VP/OT-type peptides.

Interestingly, asterotocin is the first VP/OT-type peptide with an acidic residue in position 8 of the mature neuropeptide, as opposed to a basic or neutral residue (**Fig. 3.1; Fig. 3.3**). Interestingly, the amino acid residue in position 8 of vertebrate VP (CYFQNCPRG-NH_2) and OT (CYIQNCPLG-NH_2) is the amino acid residue in which both peptides are structurally and functionally distinguished through binding to their respective receptors. This can also be seen in the mollusc *L. stagnalis*, where Lys-conopressin (CFIRNCPKG-NH_2) activates two VP/OT-type receptors with differential expression patterns in the brain and peripheral reproductive structures – *L. stagnalis* conopressin receptor 1 (LSCPR1) and LSCPR2 (van Kesteren and Geraerts, 1998; van Kesteren et al., 1995; van Kesteren et al., 1996). Interestingly, LSCPR2 is maximally activated by Lys-conopressin and the OT-like analogue Ile-conopressin (differing only via the amino acid residue at position 8), whilst LSCPR1 is only maximally activated by Lys-conopressin with Ile-conopressin approximately ten-times less potent (van Kesteren et al., 1995). However, to date, only one candidate VP/OT-type receptor has been identified in the starfish *A. rubens* (**Fig. 3.4; Fig. A5L; Fig. S4**) and in other echinoderms, including the sea urchin *S. purpuratus* in line with the presence of one VP/OT-type peptide (Elphick and Rowe, 2009). The functional significance of the acidic residue in position 8 of the mature neuropeptide remains to be explored.

3.4.2. Cardiac stomach eversion and retraction in *A. rubens*

To date, a number of neuropeptidergic and non-neuropeptidergic mediators of cardiac stomach eversion and cardiac stomach retraction, crucial to the remarkable extraoral feeding process in starfish, have been identified. In terms of cardiac

stomach eversion, it has been found that the gaseous signalling molecule nitric oxide (NO) (Elphick and Melarange, 1998) and the SALMFamide neuropeptides S1 (GFNSALMF-NH_2) and S2 (SGPYSFNSGLTF-NH_2) (Elphick et al., 1991a) cause relaxation of *in vitro* cardiac stomach preparations from the *A. rubens* (Elphick and Melarange, 1998; Elphick et al., 1995; Melarange et al., 1999). Immunocytochemical studies have also revealed that NO synthase (NOS), S1 and S2 are present in the innervation of the cardiac stomach in starfish (Martínez et al., 1994; Newman et al., 1995a; Newman et al., 1995b). Together, these data indicate that NO and the SALMFamides may mediate neural control of cardiac stomach relaxation (Elphick and Melarange, 2001). Further evidence in support of the role of S1 and S2 in cardiac stomach eversion is the observation that injection of both synthetic peptides can trigger cardiac stomach eversion when injected into the perivisceral coelom of the starfish *A. rubens* (Melarange et al., 1999). In terms of cardiac stomach retraction, it has been found that the neurotransmitter acetylcholine (ACh) causes contraction of extrinsic stomach retractor strands (Basch, 1956; Burnett and Anderson, 1955) and contraction of *in vitro* cardiac stomach preparations from *A. rubens* (Elphick et al., 1995) and it is also therefore likely that ACh mediates neural control of cardiac stomach retraction *in vivo* (Elphick et al., 1995).

3.4.3. Asterotocin: a regulator of cardiac stomach eversion

Investigation into the *in vivo* effects of asterotocin revealed that it also triggers cardiac stomach eversion in the starfish *A. rubens* (**Fig. 3.14**). However, it is important to note that only a total of 17/30 animals responded to injection of asterotocin with eversion of the cardiac stomach. This could be in part due to the relative nature of access of the peptide to muscles in the cardiac stomach via injection into the perivisceral coelom in the aboral body wall of the arms proximal to

the junctions with the central disk region. Here, access of asterotocin to cardiac stomach muscle would be limited by the compact nature of the cardiac stomach prior to eversion. Investigation into the *in vitro* pharmacological effects of asterotocin reveals that it causes dose-dependent relaxation of cardiac stomach preparations at concentrations ranging from 30 pmol l^{-1} to 1 μmol l^{-1}, with a maximal efficacy at 100 nmol l^{-1} (**Fig. 3.16**). In combination, these data suggest that asterotocin may mediate neural control of cardiac stomach eversion during the process of feeding in starfish.

It is of interest to compare the effects of asterotocin to the SALMFamides S1 and S2 in causing cardiac stomach relaxation *in vitro* and eversion *in vivo*. *In vitro*, the maximal effect of asterotocin was observed at 100 nmol l^{-1} compared to the maximal effects of the SALMFamides S1 and S2 at 10 μmol l^{-1} (Melarange et al., 1999). Moreover, asterotocin caused cardiac stomach relaxation at 30 pmol l^{-1} compared to the SALMFamides S1 at 10 nmol l^{-1} and S2 at 1 nmol l^{-1} respectively (Melarange et al., 1999). *In vivo*, asterotocin triggered cardiac stomach eversion upon injection of 10 μl of 10 μmol l^{-1} peptide compared to injection of 100 μl of 1 mmol l^{-1} S1 and S2 (Melarange et al., 1999). However, it is important to note that such a comparison between asterotocin and the SALMFamides S1 and S2 may not be physiologically relevant. Both S1 and S2 in *A. rubens* are derived from distinct precursor proteins that comprise six and seven other putative SALMFamides, respectively (see **chapter 2.3.1**). It would therefore be necessary to compare the effect of asterotocin with "cocktails" of S1 and S2 precursor-derived SALMFamides. It appears that there may be redundancy of function with respect to neuropeptidergic control of cardiac stomach eversion during starfish feeding, with both asterotocin and the SALMFamides S1 and S2 implicated in the process (Melarange et al., 1999). Such redundancy of function could be in part due to the utilisation of distinct

peptidergic systems in different contexts of the feeding process.

In combination with the effects of asterotocin *in vivo* and *in vitro*, the expression profile of asterotocin mRNA provides further evidence for the role of asterotocin in mediating neural control of cardiac stomach eversion. Importantly, asterotocin mRNA is expressed in the cardiac stomach of the starfish *A. rubens*, with cell bodies concentrated in the basiepithelial nerve plexus, with fewer cell bodies associated with the mucosal layer and visceral muscle layer (**Fig. 3.11**). This expression profile suggests that asterotocin may mediate cardiac stomach eversion by two potential mechanisms. Firstly, asterotocin may be released from processes of asterotocin-expressing cells located in the basiepithelial nerve plexus and diffuses across a thin connective tissue layer to reach the visceral muscle layer (**Fig. 3.10**). Alternatively, processes of asterotocin-expressing cells may project across the connective tissue layer to reach the visceral muscle layer directly. In this regard, developing antibodies to asterotocin and performing immunocytochemistry (ICC) will be crucial in providing insights into the functional mechanism of the peptide.

Interestingly, it should be noted that cell bodies appear to be more densely concentrated in the aboral region as opposed to the oral region of the cardiac stomach, which is densely packed with granular secretory cells (**Fig. 3.11**). It should also be noted that asterotocin mRNA was also detected in the pyloric stomach and pyloric ducts, with cell bodies sparsely located in the basiepithelial nerve plexus and mucosal layer but absent from the visceral muscle layer (**Fig. 3.12**). These results are consistent with immunoreactivity of the SALMFamide S1 in both the cardiac and pyloric stomach (Moore and Thorndyke, 1993). However, S1 expression is also associated with other parts of the digestive system including the mouth, oesophagus and pyloric caecae (Moore and Thorndyke, 1993). Taken together, the expression profile of asterotocin mRNA throughout the digestive system reinforces the role of

asterotocin in regulating cardiac stomach eversion as opposed to digestive processes.

It will be important to determine where asterotocin and, upon pharmacological characterisation, the candidate asterotocin receptor (**Fig. 3.4; Fig. A5L; Fig. S4**) are expressed in relation to one another in the context of the starfish body plan to further explore the role of asterotocin in cardiac stomach eversion. In addition to utilising mRNA ISH, it will also be important to develop antibodies to asterotocin and the candidate asterotocin receptor to perform ICC, which will help to distinguish the specific cell types that asterotocin is expressed in with regards to the mucosal layer of the gut, which includes ciliated epithelial cells (which propel food particles from the cardiac stomach through to the pyloric stomach) in addition to mucous goblet cells and granular secretory cells (Moore and Thorndyke, 1993). It will also be important to determine the signalling cascades that are involved in eliciting the cardiac stomach eversion response in order to provide insights into asterotocins mode of action. The expression of asterotocin mRNA in the cardiac stomach (and absence from retractor strands) suggests that asterotocin may exert its effects on the cardiac stomach either directly or indirectly (via cell bodies in the basiepithelial nerve plexus) by binding to its candidate receptor, which may lead to changes in secondary messenger levels or ion channel conductance (Elphick and Melarange, 2001). However, there remains the possibility that asterotocin acts indirectly by promoting the release of NO from neurons that innervate the visceral muscle layer (Elphick and Melarange, 2001). Previous studies have shown the gaseous neuronal signalling molecule NO is also involved in the process of cardiac stomach relaxation through the cyclic guanosine monophosphate (cGMP)-generating enzyme soluble guanylyl cyclase (Elphick and Melarange, 1998). However, experiments utilising 1H-[1,2,4]oxadiazolo[4,3-a]quinoxalin-1-one (ODQ), a soluble guanylyl cyclase inhibitor, on S2-induced relaxation of the cardiac stomach showed

that S2 does not function via the release of NO from neuronal cells in the cardiac stomach (Melarange et al., 1999). However, the VP/OT-type neuropeptide signalling system is coupled to the $G_q/_{11}$-phospholipase-C (PLC) signalling transduction pathway, with calcium (Ca^{2+}) release able to trigger activation of neuronal and endothelial isoforms of NOS (Stoop, 2012). It would therefore be of interest to determine whether asterotocin may in fact function through the release of NO.

3.4.4. VP/OT-type neuropeptides: an evolutionarily ancient role in feeding behaviour?

In vertebrates and other mammals, the physiological roles of VP and OT have been extensively studied. VP has a well-established role in regulating vasoconstriction and osmoregulation (Henderson and Byron, 2007; McCormick and Bradshaw, 2006), whilst OT has a well-established role in regulating reproductive processes and behaviour (Gimpl and Fahrenholz, 2001). However, in addition to their roles on peripheral target tissues, both VP and OT have also been implicated in regulating various central neuromodulatory functions including maternal-infant bonding and aggression (Broad et al., 2006; Caldwell et al., 2008; Donaldson and Young, 2008) and pair bonding and attachment (Carter, 2014; Carter et al., 1995; Cho et al., 1999; Donaldson and Young, 2008; Insel and Young, 2001; Winslow et al., 1993; Young and Wang, 2004). Interestingly, a number of functional studies in phyla and species outside of the vertebrates have revealed that the osmoregulatory, metabolic and reproductive roles of VP/OT-type peptides may be evolutionarily conserved.

The discovery that asterotocin is involved in the process of cardiac stomach eversion in *A. rubens* provides the first supporting evidence from a deuterostomian invertebrate species for an evolutionarily ancient role of VP/OT-type peptides in feeding behaviour. Previous functional studies have revealed that VP/OT-type

peptides are involved in regulating metabolic function in the mollusc *L. stagnalis* through the regulation of glycogen metabolism (Geraerts, 1992; van Kesteren et al., 1995) and in the nematode *C. elegans* through the neuromodulation of experience driven salt chemotaxis, a form of gustatory associative learning (Beets et al., 2012; Beets et al., 2013). Therefore, the role of VP/OT-type peptides as regulators of feeding behaviour can be traced back to the common ancestor of the Bilateria, with supporting data from both the deuterostomes (e.g. *A. rubens*) and the protostomes (e.g. *L. stagnalis* and *C. elegans*). However, it will be crucial to continue to investigate the roles of VP/OT-type peptides across a wider range of phyla (e.g. the cephalochordates and hemichordates) before drawing definitive conclusions on the potential evolutionarily ancient roles of the VP/OT-type neuropeptide family.

3.4.5. VP/OT-type neuropeptides: an evolutionarily ancient role in osmoregulation and reproduction?

The discovery of asterotocin has opened up the possibility to investigate the range of functions that the peptide may cover in the starfish *A. rubens*. Previous functional studies on phyla and species across the animal kingdom have revealed that VP/OT-type peptides may have an evolutionarily ancient role in both osmoregulation and reproduction, mirroring the roles of vertebrate VP and OT respectively.

In regards to the osmoregulatory role of VP, this can be observed for VP/OT-type peptides in the urochordate *S. plicata* through contraction of inhalant and exhalant siphon preparations (Ukena et al., 2008), in the echinoderm species *E. esculentus* through contraction of tube foot preparations (Elphick and Rowe, 2009), in the annelid *W. pigra* through changes body weight as a measure of water balance (Fujino et al., 1999; Oumi et al., 1996) and in the arthropod *T. castaneum* through the control of diuretic activity acting indirectly on Malpighian tubules (Aikins et al.,

2008). In support of an osmoregulatory role for asterotocin, asterotocin mRNA was detected in the nerve plexus of the adhesive region of the tube feet and was also present in marginal nerve, which innervate the outer row of tube feet in the starfish *A. rubens* (**Fig. 3.9**). Further evidence for the role of asterotocin in tube foot activity comes from the observation that injection of asterotocin into the perivisceral coelom to elicit cardiac stomach eversion leads to the elongation of tube feet in some animals (**Fig. 3.15**). Interestingly, however, initial functional studies on *in vitro* tube foot preparations have revealed that asterotocin has no effect on tube foot myoactivity (Martynuk and Elphick, unpublished data).

In regards to the reproductive role of OT, this can be observed for VP/OT-type peptides in the mollusc *L. stagnalis* through the contraction of male copulatory organs (Geraerts, 1992; van Kesteren et al., 1995), in the annelid *W. pigra* through the induction of egg-laying-like behaviours (Fujino et al., 1999; Oumi et al., 1996) and in the nematode *C. elegans* through the neuromodulation of male mating circuits (Beets et al., 2012; Beets et al., 2013). In the starfish species *A. pectinifera*, the neuropeptide GSS has been shown to induce oocyte maturation and ovulation (Mita et al., 2009). However, functional studies have also revealed that the SALMFamide S1 is involved in oocyte maturation and regulates GSS secretion from the radial nerves of the starfish species *A. pectinifera* (Mita et al., 2004). It would therefore be of interest to develop functional assays to determine whether asterotocin, like S1, has a role in regulating reproductive functions in the starfish *A. rubens*.

3.4.6. Asterotocin: a VP/OT-type peptide that causes muscle relaxation

The relaxing-type effect of asterotocin contrasts with previously reported contractile-type effects of VP/OT-type peptides in other animals found throughout the animal kingdom. For example, in other deuterostomian invertebrates, SOP causes

contraction of *in vitro* inhalant and exhalant siphon preparations from the sea squirt *S. plicata* (Ukena et al., 2008) whilst echinotocin causes contraction of *in vitro* tube foot and oesophagus preparations from the sea urchin species *E. esculentus* (Elphick and Rowe, 2009). However, the effect of asterotocin on cardiac stomach retraction is not the only relaxing-type effect of the peptide in the starfish *A. rubens*. Initial functional studies have revealed that asterotocin also causes relaxation of *in vitro* apical muscle preparations from *A. rubens* (Martynyuk and Elphick, unpublished data). However, asterotocin mRNA was not detected in the apical muscle, which suggests that it may exert its effects on apical muscle relaxation indirectly.

It is possible, however, that asterotocin may also have contractile-like effects in the starfish *A. rubens*. In this regard, injection of asterotocin into the perivisceral coelom to elicit cardiac stomach eversion also induced a number of postural changes including the thinning and curling of arms in some animals (**Fig. 3.15**). In support of a role in these contractile-like activities, asterotocin mRNA was detected in cell bodies in the epidermal layer of the body wall (**Fig. 3.13A**). It will therefore be important to determine if asterotocin causes contraction of *in vitro* body wall preparations. Moreover, asterotocin mRNA was also detected in cell bodies in the epidermal layer of the pedicellariae, which line both inside and outside of the jaw (**Fig. 3.13B**). It will therefore also be of interest to determine whether asterotocin has either a relaxing or contractile-like role on pedicellariae activity in the starfish *A. rubens*.

4. Discovery and functional characterisation of a novel neurophysin-associated neuropeptide ("NGFFYamide") in the starfish *A. rubens*

4.1. Introduction

Neurophysins are highly conserved pro-hormone derived polypeptides required for sorting the neurohypophyseal hormones vasopressin (VP) and oxytocin (OT) into the regulated secretory pathway (RSP) (Cool et al., 2008; de Bree, 2000; de Bree and Burbach, 1998). The importance of the neurophysin domain can be highlighted through the disease familial neurohypophyseal diabetes insipidus (FNDI), a water homeostasis disorder characterised by polyuria (the excretion of large volumes of diluted urine) and polydipsia (increased fluid intake). FNDI is caused by mutations in the neurophysin domain of human VP, leading to an impairment of VP biosynthesis (Babey et al., 2011). However, in addition to vertebrate VP and OT, neurophysins are associated with VP/OT-type neuropeptide precursors across the animal kingdom and thus the neurophysin domain has an ancestry that dates back to at least as far as the common ancestor of the Bilateria (Jekely, 2013; Mirabeau and Joly, 2013; Stafflinger et al., 2008; van Kesteren et al., 1995). Until recently, neurophysins were thought to be uniquely associated with VP/OT-type neuropeptide precursors. However, the discovery of a neuropeptide precursor in the sea urchin *Strongylocentrotus purpuratus* comprising myoactive peptides unrelated to VP/OT-type peptides has changed this textbook perspective (Elphick and Rowe, 2009). This neuropeptide precursor gives rise to two copies of the neuropeptide NGFFFamide, which belongs to a family of neuropeptides called the NG peptides (Elphick, 2010).

4.1.1. The NG peptide family

The NG peptides are a novel family of neuropeptides characterised by an asparagine-glycine (NG) motif that have recently been identified across the deuterostomian invertebrates (**Fig. 4.1**) (Elphick, 2010). The first NG peptide precursor to be discovered was the NGFFFamide precursor in the sea urchin *S. purpuratus*, which comprises two copies of the myoactive peptide NGFFFamide associated with a C-terminal neurophysin domain (Elphick and Rowe, 2009). Subsequently, NG peptide precursors have been identified in a number of other deuterostomian invertebrates. These include the SFRNGVamide precursor in the cephalochordate *Branchiostoma floridae* (Elphick, 2010), the NGFWNamide/NGFYNamide precursor in the hemichordate *Saccoglossus kowalevskii* (Elphick, 2010) and the NGIWYamide precursor in the sea cucumber *Apostichopus japonicus* (Elphick, 2012). Interestingly, however, the NGIWYamide precursor in the sea cucumber *A. japonicus* is derived from a precursor protein that lacks a neurophysin domain, which is likely to have been lost in the holothurian lineage (Elphick, 2012).

(A)

(B)

B.flo (Chordata)	SFRNGV-NH_2(x2)
S.kow (Hemichordata)	NGFWN-NH_2(x5)& NGFYN-NH_2 (X1)
S.pur (Echinodermata)	NGFFF-NH_2(x2)
A.jap (Echinodermata)	NGIWY-NH_2(x5)

Fig. 4.1. The NG peptide family. (A) Schematic depicting the structural organisation of NG peptide precursors in the deuterostomian invertebrates. Signal peptides are represented in blue, putative neuropeptides are represented in red, dibasic cleavage sites are represented in green and C-terminal neurophysin domains are represented in purple. **(B)** Alignments of putative neuropeptides derived from NG peptide precursors. Conserved NG motifs highlighted in yellow and the number of copies of each peptide is represented in parentheses. *A.jap*, *Apostichopus japonicus; B.flo*, *Branchiostoma floridae*; *S.kow*, *Saccoglossus kowalevskii*; *S.pur*, *Strongylocentrotus purpuratus*.

4.1.2. Functional roles of the NG peptides

Sequencing of the sea urchin genome (Sodergren et al., 2006) led to the discovery of the first NG peptide precursor in the deuterostomian invertebrates – the NGFFFamide precursor (Elphick and Rowe, 2009). Subsequently, functional studies have revealed that NGFFFamide causes contraction of *in vitro* tube foot and oesophagus preparations from the sea urchin species *Echinus esculentus* (Elphick and Rowe, 2009). However, the first NG peptide to be identified biochemically was NGIWYamide in the sea cucumber *A. japonicus* (Iwakoshi et al., 1995). NGIWYamide was identified through a bioassay utilising radial longitudinal muscle (RLM) and intestinal muscle preparations to screen for myoactive peptides in *A. japonicus* (Iwakoshi et al., 1995; Ohtani et al., 1999). Two SALMFamide neuropeptides were identified as muscle relaxants on intestinal muscle preparations – GYSPFMFamide and FKSPFMFamide (Iwakoshi et al., 1995; Ohtani et al., 1999). Moreover, one neuropeptide that causes muscle contraction in RLM and intestinal muscle preparations was also identified – this was NGIWYamide (Iwakoshi et al., 1995; Ohtani et al., 1999). Subsequently, NGIWYamide has been shown to control the stiffness of connective tissue in the body wall dermis (Birenheide et al., 1998) and cause contraction of tentacle preparations and inhibition of the rhythmic contractile activity of intestine preparations (Inoue et al., 1999). Accordingly, immunocytochemical studies have revealed that NGIWYamide has a widespread distribution in neurons innervating the tube feet, tentacles, intestine and body wall dermis, consistent with its effects *in vitro* (Inoue et al., 1999). Recently, it has also been shown that NGIWYamide induces oocyte maturation and gamete spawning in the sea cucumber *A. japonicus* (Kato et al., 2009). Interestingly, NGIWYamide has also been shown to cause contraction of *in vitro* tube foot preparations from the starfish species *Asterina pectinifera* (Saha et al., 2006). Accordingly,

immunocytochemical studies have detected NGIWYamide-like immunoreactivity in neurons innervating the tube feet (Saha et al., 2006). However, until now, the identity of the NGIWYamide-like peptide in *A. pectinifera* was unknown. The discovery of an NG peptide ("NGFFYamide") in the starfish *A. rubens* (see **chapter 2.3.2.2**) has therefore provided further opportunities to investigate the function of the NG peptides in an echinoderm species.

4.1.3. Investigating the NG peptides in *A. rubens*

Functional studies have revealed that the processes of cardiac stomach eversion and cardiac stomach retraction in the common European starfish *Asterias rubens* is at least in part controlled by a number of neuropeptidergic and non-peptidergic mediators. Cardiac stomach relaxation and eversion is triggered by the gaseous signalling molecule nitric oxide (NO) (Elphick and Melarange, 1998), the SALMFamides S1 and S2 (Elphick et al., 1995; Melarange et al., 1999) and a VP/OT-type peptide ("asterotocin") (see **chapter 3**). On the other hand, it has been hypothesised that cardiac stomach retraction is controlled by the neurotransmitter acetylcholine (ACh), which has been shown to cause contraction of *in vitro* cardiac stomach preparations from the starfish *A. rubens* (Elphick et al., 1995). Therefore, with both neuropeptidergic (S1, S2 and asterotocin) and non-peptidergic (NO) control of cardiac stomach eversion, it is likely that there are also neuropeptidergic mediators acting in parallel with ACh to mediate the process of cardiac stomach retraction in the starfish *A. rubens*. To date, functional studies in the sea cucumber *A. japonicus* (Kato et al., 2009) and the sea urchin species *E. esculentus* (Elphick and Rowe, 2009) have revealed that the NG peptides act as muscle contractants in the echinoderms. Therefore, the discovery of NGFFYamide (see **chapter 2.3.2.2**) has provided an opportunity to investigate a potential role of this neuropeptide as a

mediator of cardiac stomach contraction in the starfish *A. rubens*.

4.1.4. Aims and objectives

Herein, the analysis of neural transcriptome data from the starfish *A. rubens* has led to the discovery of a precursor protein that gives rise to two copies of the putative neuropeptide NGFFYamide (see **chapter 2.3.2.2**). This discovery of NGFFYamide has provided an opportunity to:

(1) Confirm the sequence of the NGFFYamide precursor by cDNA cloning and sequencing.

(2) Determine the structure of the NGFFYamide peptide through the use of mass spectrometric methods.

(3) Investigate the expression of the NGFFYamide precursor in the starfish *A. rubens* using mRNA *in situ* hybridisation (mRNA ISH).

(4) Investigate the physiological functions of NGFFYamide in the starfish *A. rubens* using *in vitro* and *in vivo* pharmacological techniques.

4.2. Methods

4.2.1. Bioinformatic identification of the NGFFYamide precursor in *A. rubens*

To search for a transcript encoding an NG peptide precursor in the starfish *A. rubens*, the sea urchin *S. purpuratus* NG peptide precursor ("NGFFFamide" precursor) protein sequence (Elphick and Rowe, 2009) was submitted as a query in a tBLASTn search of the *A. rubens* neural transcriptome using SequenceServer software (http://www.sequenceserver.com/) (Priyam et al., in prep). The protein sequence encoded by the top-hit (expected (e)-value) transcript was then analysed for the presence of an N-terminal signal peptide as predicted by SignalP 3.0 (Bendtsen et al., 2004), putative NG peptides flanked by dibasic cleavage sites and a C-terminal neurophysin domain, which is characterised by the presence of fourteen cysteine (C) residues (Elphick, 2010; Elphick and Rowe, 2009). Subsequently, the protein sequence was submitted in a tBLASTn search of The National Centre for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine sequence similarity to known neuropeptide or peptide hormone precursors.

4.2.2. Cloning and sequencing of NGFFYamide precursor

Total RNA generated from the same specimen of *A. rubens* used for Illumina Hi-Seq (see **chapter 2.2.2**) was used to generate cDNA using the Quantitect Reverse Transcription Kit (QIAgen). NGFFYamide precursor cDNA containing the entire coding region was amplified by PCR using Phusion High-Fidelity PCR Master Mix (NEB) and the oligos: 5'-AGACCTTATAGGCTTAGAG-3' and 5'-GTCATCACGATAACACTC-3'. The PCR product was gel-extracted and purified

using the QIAquick Gel Extraction Kit (QIAgen) before being blunt-end cloned into a pBluescript SKII (+) vector (Agilent Technologies) cut with the EcoRV-HF (NEB) restriction endonuclease. Three clones were sequenced from both the T3 and T7 sequencing primer sites for sequence confirmation (Eurofins). A plasmid maxiprep was subsequently performed using the QIAgen Plasmid Maxi Kit (QIAgen).

4.2.3. Mass spectrometric confirmation of NGFFYamide

Radial nerve cords were dissected from five adult specimens of *A. rubens* using a method described previously (Chaet, 1964). Neuropeptides were then extracted in 1 ml of 80 % acetone on ice (Elphick et al., 1991a). Acetone was subsequently removed by evaporation using nitrogen, with the aqueous fraction centrifuged at 11,300 g in a MiniSpin® microcentrifuge (Eppendorf) for 10 minutes and the remaining supernatant stored at −80 °C. Prior to mass spectrometry (MS), the extract was thawed and filtered through a 0.22 μm Costar® Spin-X® centrifuge tube filter (Sigma-Aldrich) to remove particulates. The extract was then analysed by nanoflow liquid chromatography (LC) with electrospray ionisation (ESI) quadrupole time-of-flight tandem MS (nano LC-ESI-MS/MS) using a nanoAcquity ultra performance LC (UPLC) system coupled to a Q-TOF Ultima Global Mass Spectrometer™ (Waters Corporation) and MassLynx v.4.0 software (Waters Corporation). The mobile phases used for the chromatographic separation were 0.1 % aqueous formic acid (termed mobile phase A) and 0.1 % formic acid in acetonitrile (termed mobile phase B). An aliquot containing 5 μl of the *A. rubens* radial nerve cord extract was applied to a Symmetry C18® trapping column (180 μm x 20 mm, 5 μm particle size, 100 Å pore size) (Waters Corporation) using 99.9 % mobile phase A at a flow rate of 15 μl min^{-1} for 1 minute, after which the fluidic flow path included the HSS T3 analytical capillary column (75 μm x 150 mm, 1.8 μm particle size, 100 Å pore size) (Waters

Corporation). A linear gradient of 5–40 % mobile phase B over 45 minutes was utilised with a total run time of 60 minutes. The nanoflow ESI source conditions utilised a 3.5 kV capillary voltage, 25 V sample cone voltage and an 80 °C source temperature. A data-dependent acquisition was performed that would trigger a MS/MS scan on any singly charged peptide having a mass/charge (m/z) ratio of 646.2989, or a doubly charged peptide of m/z of 323.6534, with a tolerance of 150 mDa allowed on the precursor m/z. MS/MS spectra, obtained from data-dependent acquisition, was processed using MassLynx v.4.0 software (Waters Corporation). Spectra were combined and processed using the MaxEnt 3 algorithm to generate singly charged, monoisotopic spectra for interpretation and manual validation.

4.2.4. Localisation of the NGFFYamide precursor in *A. rubens* using mRNA *in situ* hybridisation (mRNA ISH)

The mRNA *in situ* hybridisation (mRNA ISH) methodology (probe synthesis, tissue fixation, tissue sectioning, probe hybridisation and immunodetection) was performed in accordance to the methodology utilised for the asterotocin precursor (see **chapter 3.2.4**).

4.2.5. Analysis of the *in vitro* effects of NGFFYamide on cardiac stomach preparations from *A. rubens*

Cardiac stomachs were dissected from specimens of *A. rubens* and set up in a 20 ml organ bath maintained at 11 °C, as described previously (**Fig. 3.2**) (Elphick et al., 1995; Melarange et al., 1999). However, unlike for analysis of the *in vitro* effects of synthetic asterotocin (see **chapter 3.2.6**), the organ bath contained artificial seawater that was not supplemented with 30 mM potassium chloride (KCl). Stock solutions of synthetic NGFFYamide (NGFFY-NH_2) or NGFFFamide (NGFFF-NH_2) (Peptide Protein Research Ltd.) were prepared in autoclaved water and then added individually to the organ bath to achieve final concentrations ranging from 30 pmol l^{-1} to 1 μmol l^{-1} upon the cardiac stomach maintaining a stabilised state. Cardiac stomach contraction was recorded using an isotonic transducer (Harvard Apparatus) linked to a Goerz SE 120 chart recorder (Recorderlab). The preparation was subsequently washed with artificial seawater between different concentrations to allow the cardiac stomach preparation to return to a stabilised state.

4.2.6. Analysis of the *in vivo* effects of NGFFYamide in *A. rubens*

Informed by the pharmacological effect of synthetic NGFFYamide on cardiac stomach preparations *in vitro*, experiments were performed to investigate if NGFFYamide triggers cardiac stomach retraction *in vivo.* A total of ten specimens of *A. rubens* withheld from a food supply (mussels) for 1 week were placed in a glass tank containing 2 % magnesium chloride ($MgCl_2$) dissolved in artificial seawater. $MgCl_2$ acts as a muscle relaxant in marine invertebrates (Mayer, 1909) and reproducibly causes eversion of the cardiac stomach in *A. rubens*, typically within a period of approximately 15-30 minutes (Elphick, unpublished observations). Hamilton 75N 5 µl syringes (Sigma-Aldrich) were used to inject synthetic NGFFYamide (NGFFY-NH_2) (Peptide Protein Research Ltd.) into the perivisceral coelom of animals at two sites in the aboral body wall of the arms proximal to the junctions with the central disk region. Animals were first injected with 10 µl autoclaved water (control) and video recorded for 220 seconds. The same animals were then injected with 10 µl of 100 nmol l^{-1} synthetic NGFFYamide (concentration selected based on results from *in vitro* pharmacology) and video recorded for 4 minutes. Care was taken to inject test agents into the perivisceral coelom and not into the cardiac stomach. Static images from video recordings were captured at 20-second intervals from the time of injection. Then the two-dimensional area of everted cardiac stomach was measured from the images using ImageJ software (http://rsb.info.nih.gov/ij/) and normalised as a percentage of the area of cardiac stomach everted at the time of injection.

4.3. Results

4.3.1. Bioinformatic identification and cloning of the NGFFYamide precursor in *A. rubens*

In BLAST analysis of the *A. rubens* neural transcriptome data using the sea urchin *S. purpuratus* NGFFFamide precursor as a query, the top-hit (e-value: $2e^{-15}$) was contig 1104160 (1268 bp), which encodes a 239-residue protein termed the NGFFYamide precursor. The sequence of the NGFFYamide precursor has been confirmed by cDNA cloning and sequencing (**Fig. 4.2**). The precursor protein comprises a 23-residue N-terminal signal peptide (residues 1-23), two tandem copies of a putative NG peptide (NGFFYG) (residues 124-129 and residues 132-137), with C-terminal glycine (G) residues predicted to be subject to amidation (NGFFYamide; NGFFY-NH_2). The NGFFYamide precursor also has an 80-residue C-terminal neurophysin domain comprising fourteen conserved cysteine (C) residues, hitherto characteristic of VP/OT-type precursors (residues 153-232). The sequence of the 1268 bp NGFFYamide precursor transcript has been deposited in the GenBank database and assigned accession number KC977457.

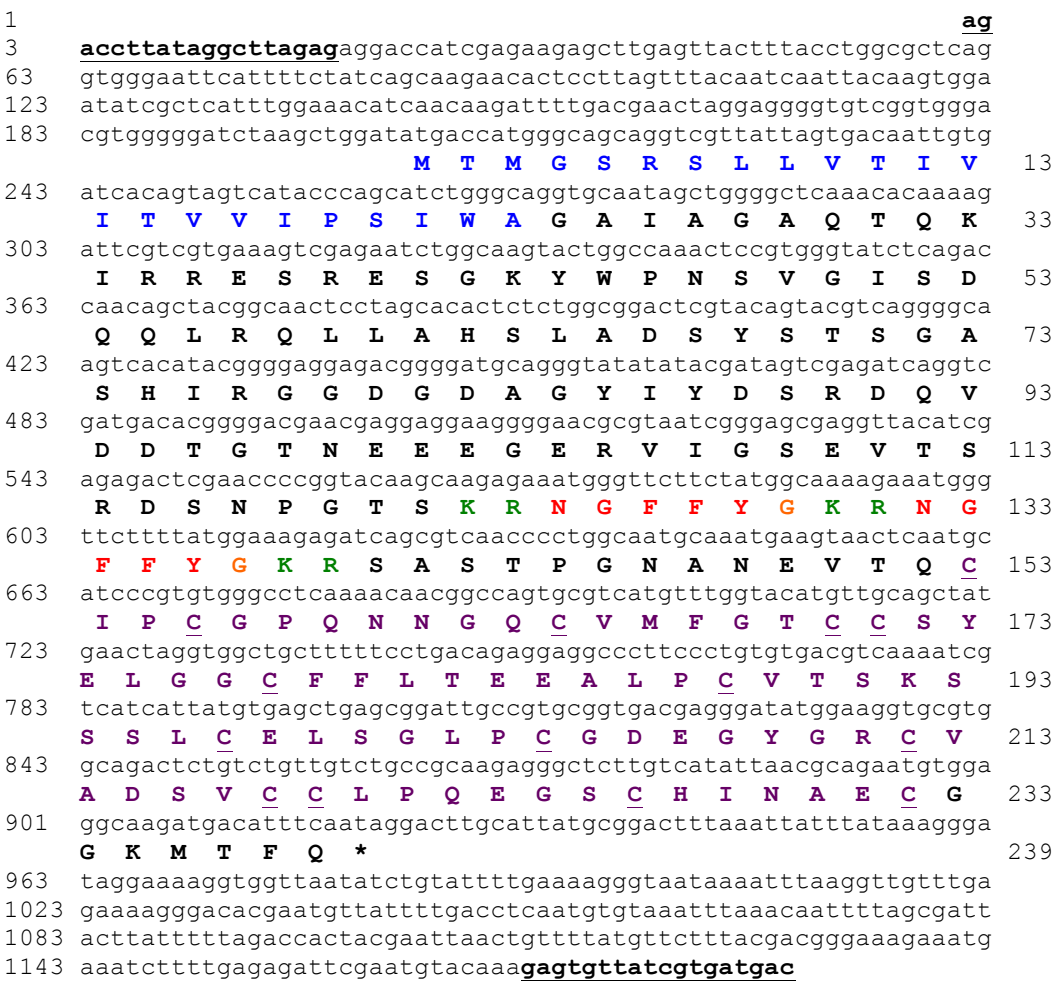
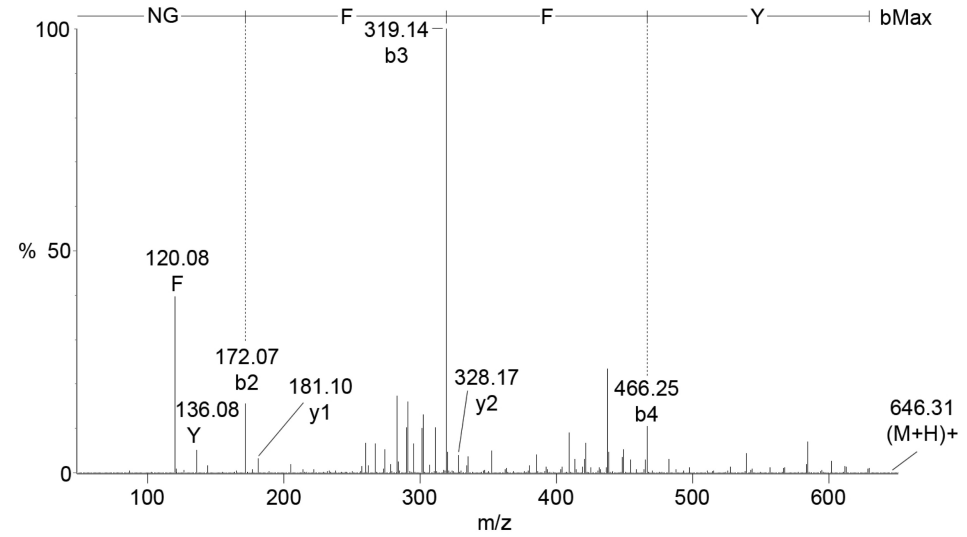


Fig. 4.2. *A. rubens* NGFFYamide precursor cDNA sequence. The nucleotide sequence (lowercase, 1189 bases) encoding the precursor protein (uppercase, 239 amino acid residues) is shown. Primers used for cloning the NGFFYamide precursor cDNA are represented in bold and underlined text. The predicted signal peptide is represented in blue, two tandem copies of the putative NGFFYamide peptide are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites (KR) are represented in green. The C-terminal region of the precursor comprises a neurophysin domain (with fourteen characteristic cysteine (C) residues underlined), which is represented in purple. The asterisk shows the position of the stop codon.

4.3.2. Mass spectrometric confirmation of NGFFYamide

In order to perform both *in vitro* and *in vivo* functional studies, it was important to confirm the sequence and structure of the putative NGFFYamide peptide. Cloning and sequencing of the NGFFYamide precursor confirmed the presence of two copies of an NG peptide ("NGFFYamide") with the sequence NGFFYG (**Fig. 4.2**). However, it is also important to confirm whether the C-terminal glycine (G) residue is converted to an amide group. Nano LC-ESI-MS/MS using synthetic NGFFYamide as a standard was first implemented, resulting in the peptide eluting at a retention time of 30.3 minutes with the singly charged species observed at a m/z of 646.3 (**Fig. 4.3A**). Subsequent analysis of the *A. rubens* radial nerve cord extract under identical conditions revealed that a single charged peptide with an m/z of 646.3 eluted at the same retention time to synthetic NGFFYamide (**Fig. 4.3B**). Both peptides were subjected to MS/MS during the experiment and the resulting deconvoluted, singly charged, monoisotopic spectra were compared, confirming the presence of NGFFYamide in the radial nerve cord extract (**Fig. 4.3A-B**).

(A)



(B)

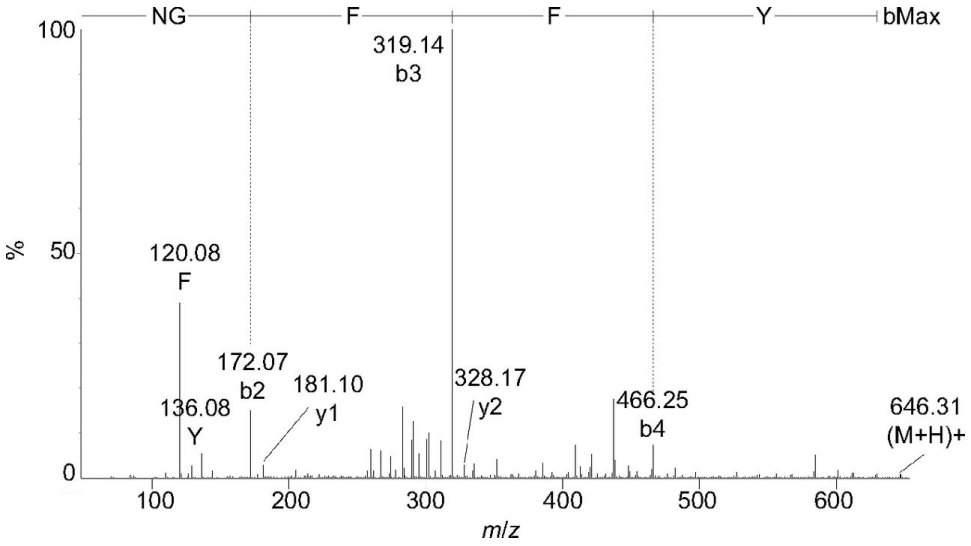


Fig. 4.3. Mass spectrometric confirmation that NGFFYamide is present in the radial nerve cord of *A. rubens*. **(A)** MS/MS spectrum of synthetic NGFFYamide peptide. **(B)** MS/MS spectrum of NGFFYamide in an acetone extract of radial nerve cords of *A. rubens*. The deconvoluted, monoisotopic, singly charged spectrum derived from MS/MS data is shown, with the b-series of fragment ions annotated (b2, b3 and b4). Also labelled are two fragment ions from the y-series (y1 and y2), immonium ions from phenylalanine (F) and tyrosine (Y) residues and the NGFFYamide peptide ion (646.31).

4.3.3. Expression profile of NGFFYamide mRNA in *A. rubens*

Serial sections were taken through the central disk and one arm of the starfish *A. rubens* to allow visualisation of the expression of NGFFYamide precursor mRNA (or "NGFFYamide mRNA" for brevity) (**Fig. 3.6**). Expression of NGFFYamide mRNA was observed in the radial nerve cord, circumoral nerve ring, marginal nerve, tube feet, cardiac stomach, pyloric stomach, pyloric ducts, the apical muscle and the coelomic epithelium, which will be discussed on a one-by-one basis. Note that expression of NGFFYamide mRNA was not detected in control experiments utilising a NGFFYamide mRNA sense probe to detect non-specific binding (an example is provided in **Fig. 4.4**).

4.3.3.1. Radial nerve cord and circumoral nerve ring

In transverse sections of the arm, expression of NGFFYamide mRNA was observed in the ectoneural region of the radial nerve cord but not in the hyponeural region (**Fig. 4.5A-B**). The labelled cell bodies were, largely, laterally concentrated at the periphery of the radial nerve cord in the epithelium of the ectoneural nerve plexus (**Fig. 4.5A-B**). NGFFYamide mRNA expression was also present in cell bodies located in the lateral branches of the ectoneural region of the radial nerve cord (**Fig. 4.5D**). In sections through the central disk, expression of NGFFYamide mRNA was also observed in the ectoneural region of the circumoral nerve ring (**Fig. 4.5C**), consistent with the expression pattern observed in the ectoneural region of the radial nerve cord.

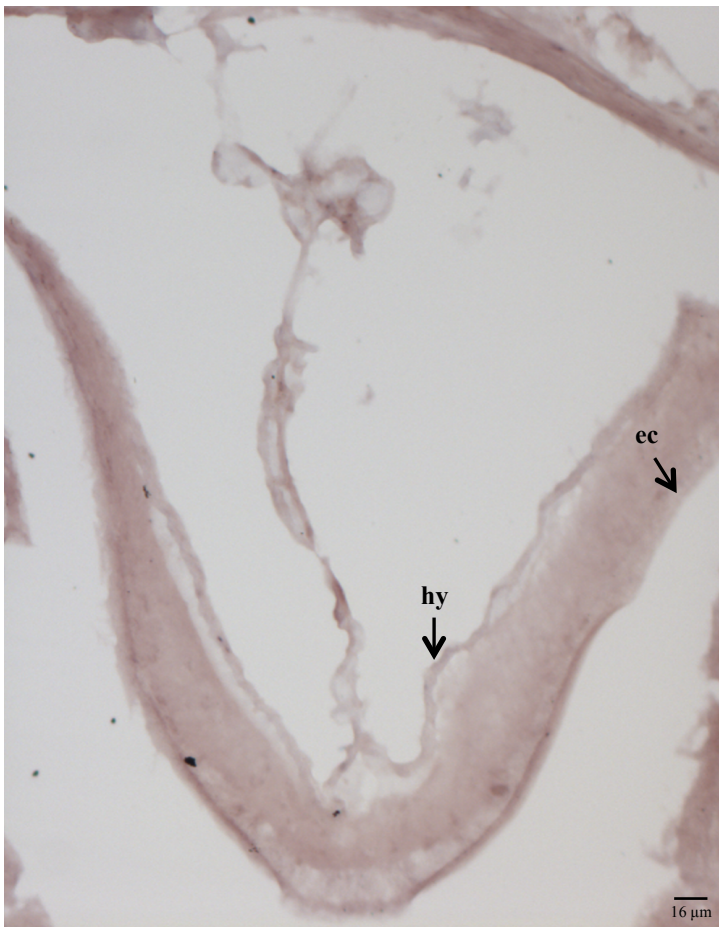


Fig. 4.4. Control experiment for NGFFYamide mRNA expression in the radial nerve cord in *A. rubens*. Transverse section through an arm of *A. rubens* displaying the radial nerve cord. NGFFYamide mRNA expression was not detected in the radial nerve cord utilising a NGFFYamide mRNA sense probe to detect non-specific binding. *ec* = *ectoneural*; *hy* = *hyponeural*.

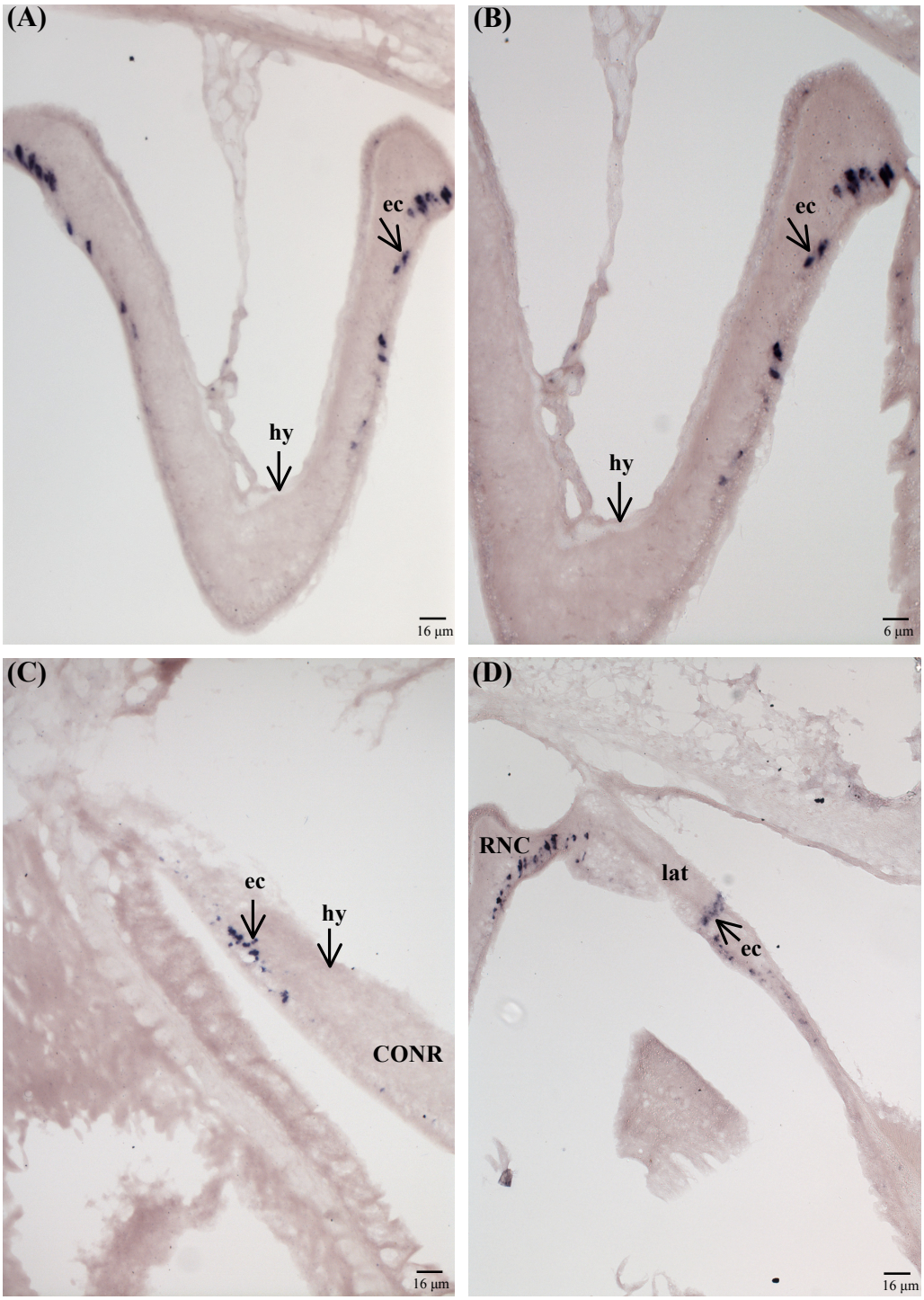

(A)
ec
hy
16 µm
(B)
ec
hy
6 µm
(C)
ec
hy
CONR
16 µm
(D)
RNC
lat
ec
16 µm

Fig. 4.5. NGFFYamide mRNA expression in the radial nerve cord, lateral branches of the radial nerve cord and circumoral nerve ring in *A. rubens*. (A-B) Transverse section through an arm of *A. rubens* displaying the radial nerve cord at two magnifications. NGFFYamide mRNA expression was detected in the epithelium of the ectoneural nerve plexus of the radial nerve cord. **(C)** Transverse section through the central disk of *A. rubens* displaying the circumoral nerve ring. NGFFYamide mRNA expression was detected in the ectoneural region of the circumoral nerve ring. **(D)** Transverse section through an arm of *A. rubens* displaying the lateral branches of the radial nerve cord. NGFFYamide mRNA expression was detected in the ectoneural region of the lateral branches of the radial nerve cord. *CONR = circumoral nerve ring*; *ec = ectoneural*; *hy = hyponeural*; *lat = lateral branch of radial nerve cord*; *RNC = radial nerve cord*.

4.3.3.2. Tube feet and marginal nerve

In transverse sections through the arm and central disk, expression of NGFFYamide mRNA was observed in cells located in the adhesive region of tube feet suckers (**Fig 4.6A-B**). Moreover, expression of NGFFYamide mRNA was also observed in the basal nerve ring (**Fig. 4.6A-B**). However, expression of NGFFYamide mRNA was not observed in the sub-epithelial nerve plexus. NGFFYamide mRNA expression was also present in the marginal nerve, which innervates the outer row of tube feet (**Fig. 4.6C-D**).

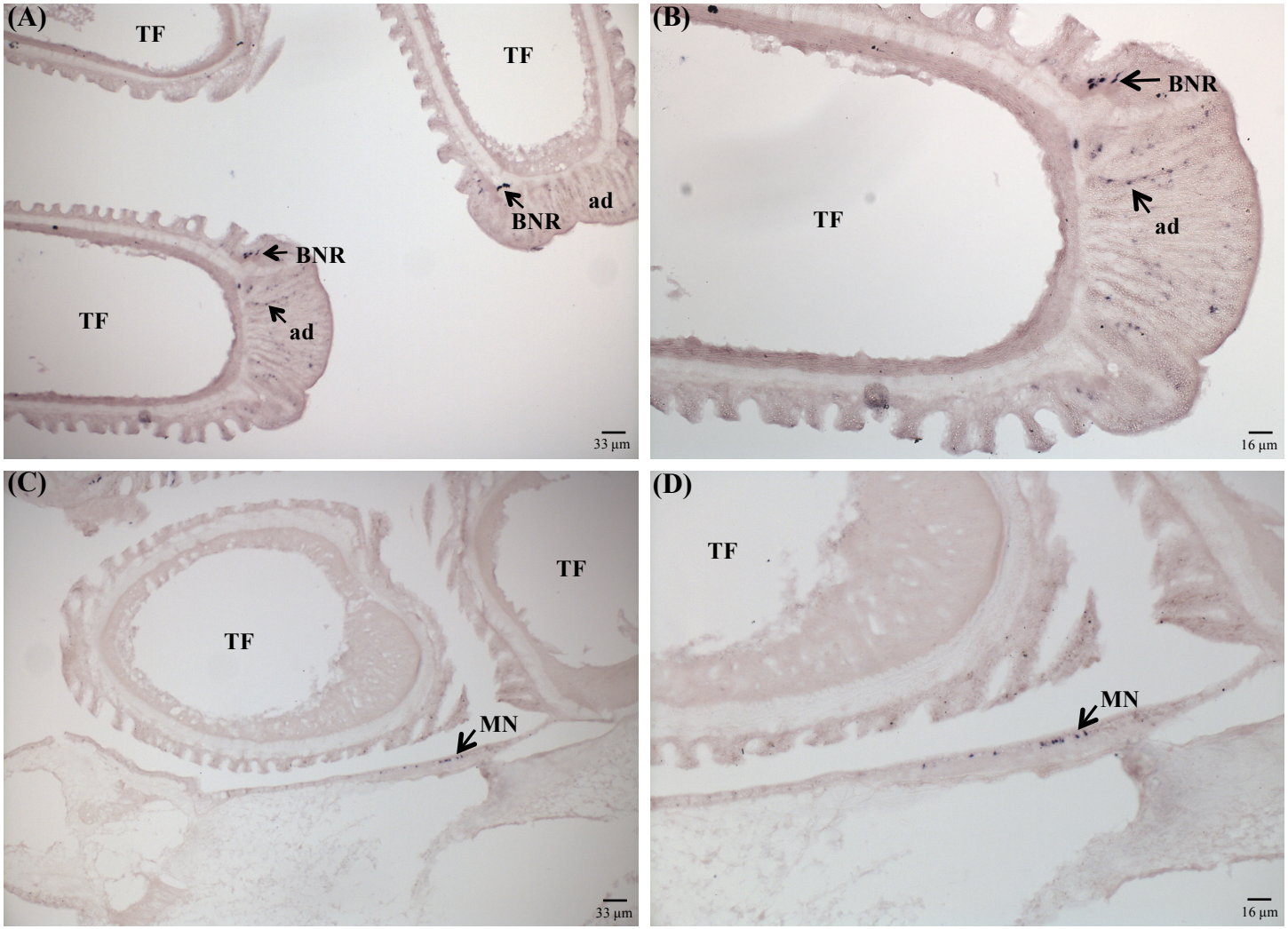

(A)
TF
TF
BNR
ad
BNR
TF
ad
33 µm
(B)
BNR
TF
ad
16 µm
(C)
TF
TF
MN
33 µm
(D)
TF
MN
16 µm

Fig. 4.6. NGFFYamide mRNA expression in the tube feet and marginal nerve in *A. rubens*. **(A-B)** Transverse section through an arm of *A. rubens* displaying tube feet at two magnifications. NGFFYamide mRNA expression was detected in the nerve plexus in the adhesive region of the tube foot sucker and in the basal nerve ring. **(C-D)** Transverse section through an arm of *A. rubens* displaying a marginal nerve supplying innervation to a tube foot at two magnifications. NGFFYamide mRNA expression was detected in the marginal nerve. *ad = adhesive region of tube foot*; *BNR = basal nerve ring*; *MN marginal nerve*; *TF = tube foot*.

4.3.3.3. Digestive system

In sections through the central disk, expression of NGFFYamide mRNA was observed in the cardiac stomach (**Fig. 4.7**) with sparser expression detected in the pyloric stomach and very sparse expression detected in the pyloric ducts (**Fig. 4.8A-B** and **Fig. 4.8D**). However, expression of NGFFYamide mRNA was not observed in the peristomial membrane, oesophagus, pyloric caecae (**Fig. 4.8C**), intestine, rectal caecae (**Fig. 4.8A**) or retractor strands. Labelled cell bodies appear to be more densely concentrated in the aboral region as opposed to the oral region of the cardiac stomach, which is densely packed with granular secretory cells (**Fig. 4.7**). In both the aboral and oral regions of the cardiac stomach, however, labelled cells bodies are largely concentrated in the basiepithelial nerve plexus (**Fig 4.7**). Labelled cell bodies are also located in the mucosal layer, with very sparse expression in the visceral muscle layer (**Fig. 4.7**). In the pyloric stomach, labelled cell bodies are sparsely located in the basiepithelial nerve plexus but are absent from the mucosal and visceral muscle layers (**Fig. 4.8A-B**). In the pyloric ducts, cell bodies are sparsely located in the basiepithelial nerve plexus and the mucosal layer but are absent from the visceral muscle layer (**Fig. 4.8D**).

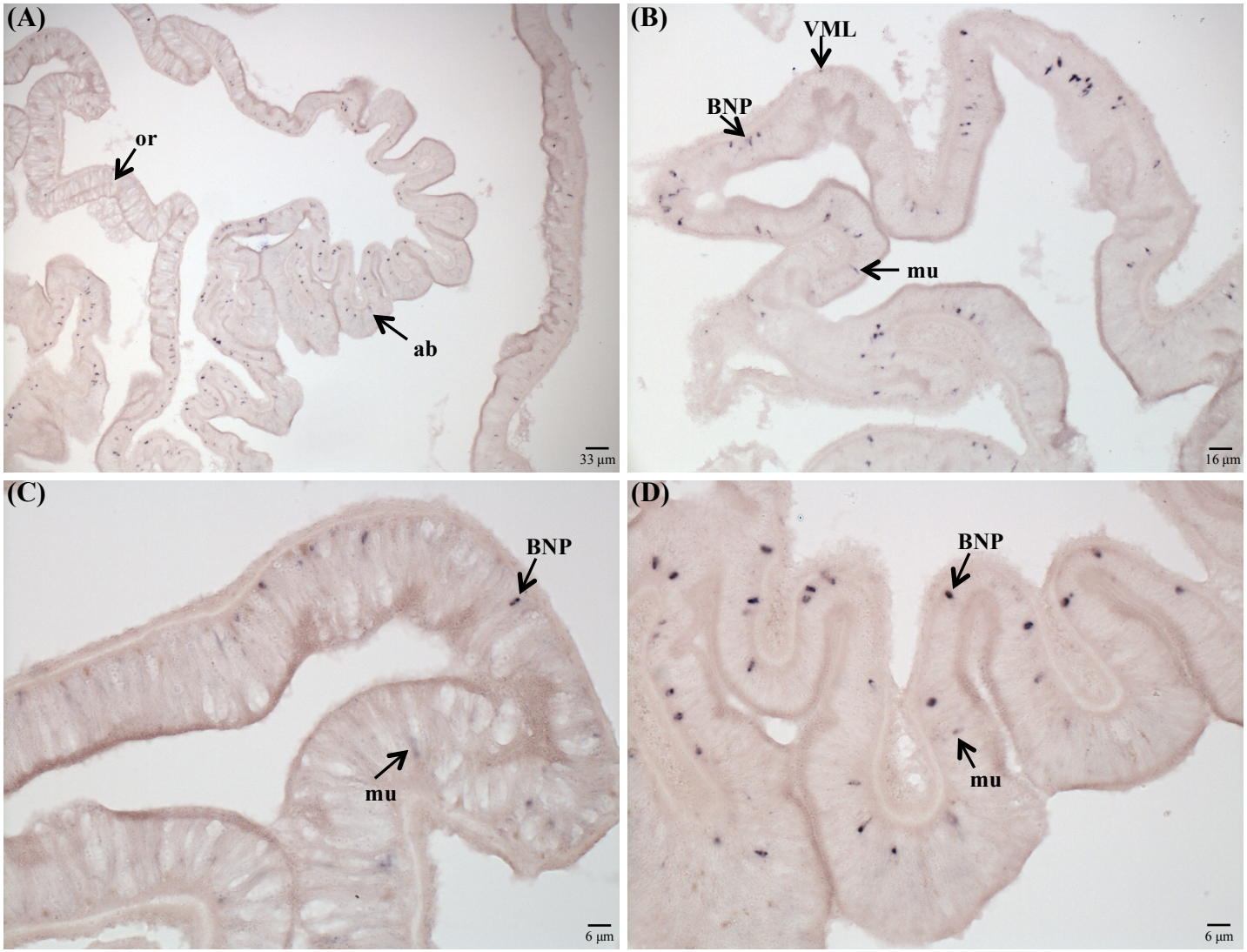

(A)
or
ab
33 µm
(B)
VML
BNP
mu
16 µm
(C)
BNP
mu
6 µm
(D)
BNP
mu
6 µm

Fig. 4.7. NGFFYamide mRNA expression in the cardiac stomach of *A. rubens*. **(A)** Transverse section through the central disk of *A. rubens* displaying the oral and aboral regions of the cardiac stomach. **(B)** Transverse section through the central disk of *A. rubens* displaying the aboral region of the cardiac stomach. NGFFYamide mRNA expression was detected in the basiepithelial nerve plexus, with sparse expression in the mucosal and visceral muscle layers. **(C)** Transverse section through the central disk of *A. rubens* displaying an oral region of the cardiac stomach. NGFFYamide mRNA expression was detected in the basiepithelial nerve plexus, with sparse expression in the mucosal layer. **(D)** Transverse section through the central disk of *A. rubens* displaying the aboral region of the cardiac stomach. NGFFYamide mRNA expression was detected in the basiepithelial nerve plexus, with sparse expression in the mucosal layer. *ab = aboral cardiac stomach; BNP = basiepithelial nerve plexus*; *mu = mucosal layer*; *or = oral cardiac stomach*; *VML = visceral muscle layer*.

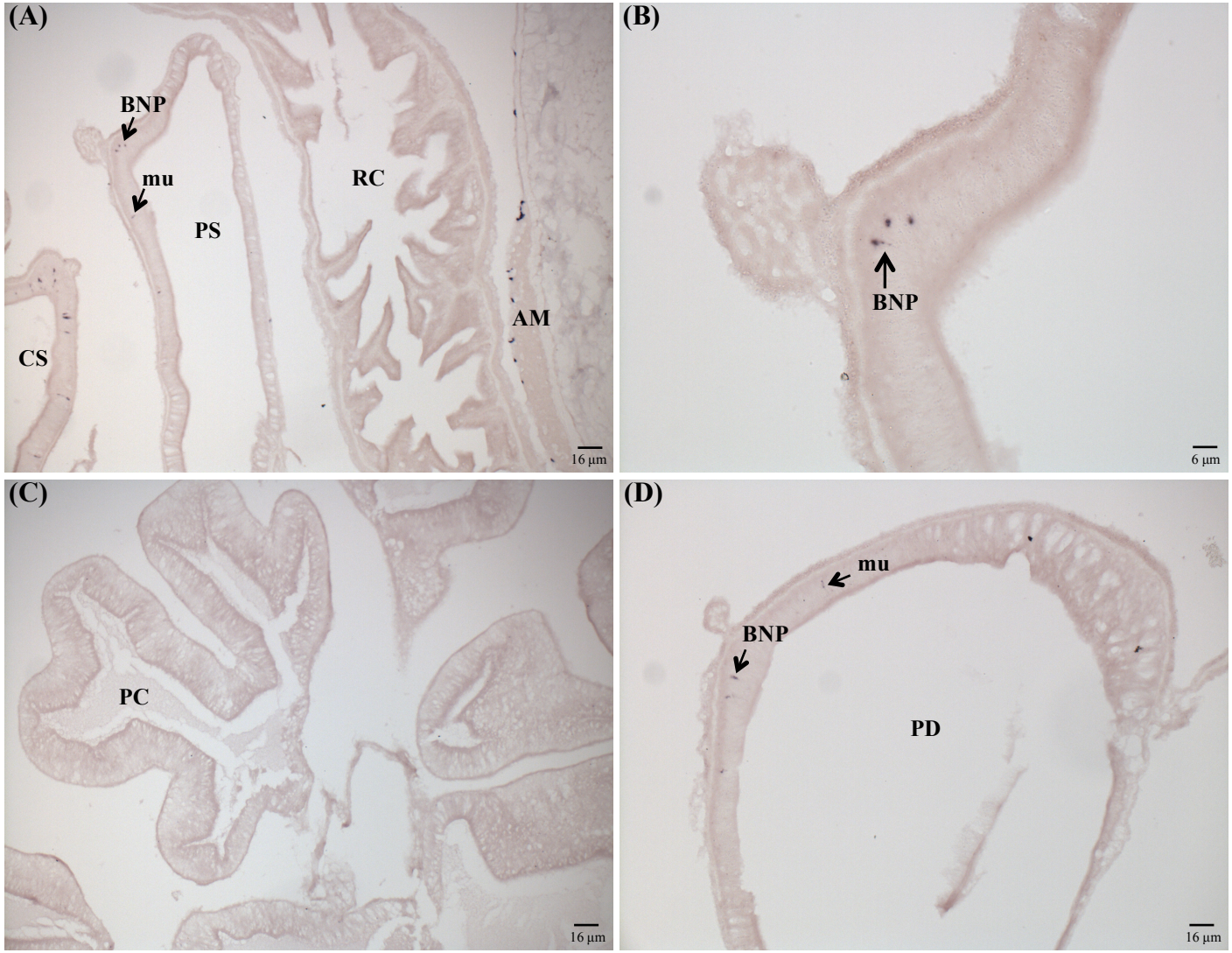

(A)
BNP
mu
RC
PS
AM
CS
16 µm
(B)
BNP
6 µm
(C)
PC
16 µm
(D)
mu
BNP
PD
16 µm

Fig. 4.8. NGFFYamide mRNA expression in regions of the digestive system in *A. rubens*. **(A)** Transverse section through the central disk of *A. rubens* displaying the pyloric stomach and rectal caecae. NGFFYamide mRNA expression was detected in the basiepithelial nerve plexus and mucosal layer of the pyloric stomach. Note that NGFFYamide mRNA expression can also be seen in the apical muscle. **(B)** Transverse section through the central disk of *A. rubens* displaying the pyloric stomach. NGFFYamide mRNA expression was detected in the basiepithelial nerve plexus and mucosal layer of the pyloric stomach. **(C)** Transverse section through the central disk of *A. rubens* displaying pyloric caecae. NGFFYamide mRNA expression was not detected in the pyloric caecae. **(D)** Transverse section through the central disk of *A. rubens* displaying a pyloric duct. NGFFYamide mRNA expression was detected in the basiepithelial nerve plexus and mucosal layer of the pyloric duct. *AM = apical muscle*; *BNP = basiepithelial nerve plexus*; *CS = cardiac stomach*; *mu = mucosal* layer; PC = *pyloric caecae*; *PD = pyloric duct*; *PS = pyloric stomach*; *RC = rectal caecae*; *VML = visceral muscle layer*.

4.3.3.4. Apical muscle and coelomic epithelium

The apical muscle of the starfish *A. rubens* is a muscular thickening located on the inside of the dorsal body wall (Moore and Thorndyke, 1993). The apical muscle is bound to a thin band of circular sub-epidermal muscle at the junction between the arm and central disk, which has been termed the aboral autotomy plane (Moore and Thorndyke, 1993). The apical muscle is also bound to a ciliated coelomic epithelial layer and an overlying nerve plexus, which is continuous with innervation of the gut (Moore and Thorndyke, 1993). The expression of NGFFYamide mRNA was observed in cell bodies in the coelomic epithelial layer lining the apical muscle, which forms the coelomic lining of the apical muscle (**Fig. 4.9A)**. The expression of NGFFYamide was also observed in cell bodies in the coelomic epithelium, located at the junction between the central disk and arm, which separates the body wall from the coelomic cavity (**Fig. 4.9B**).

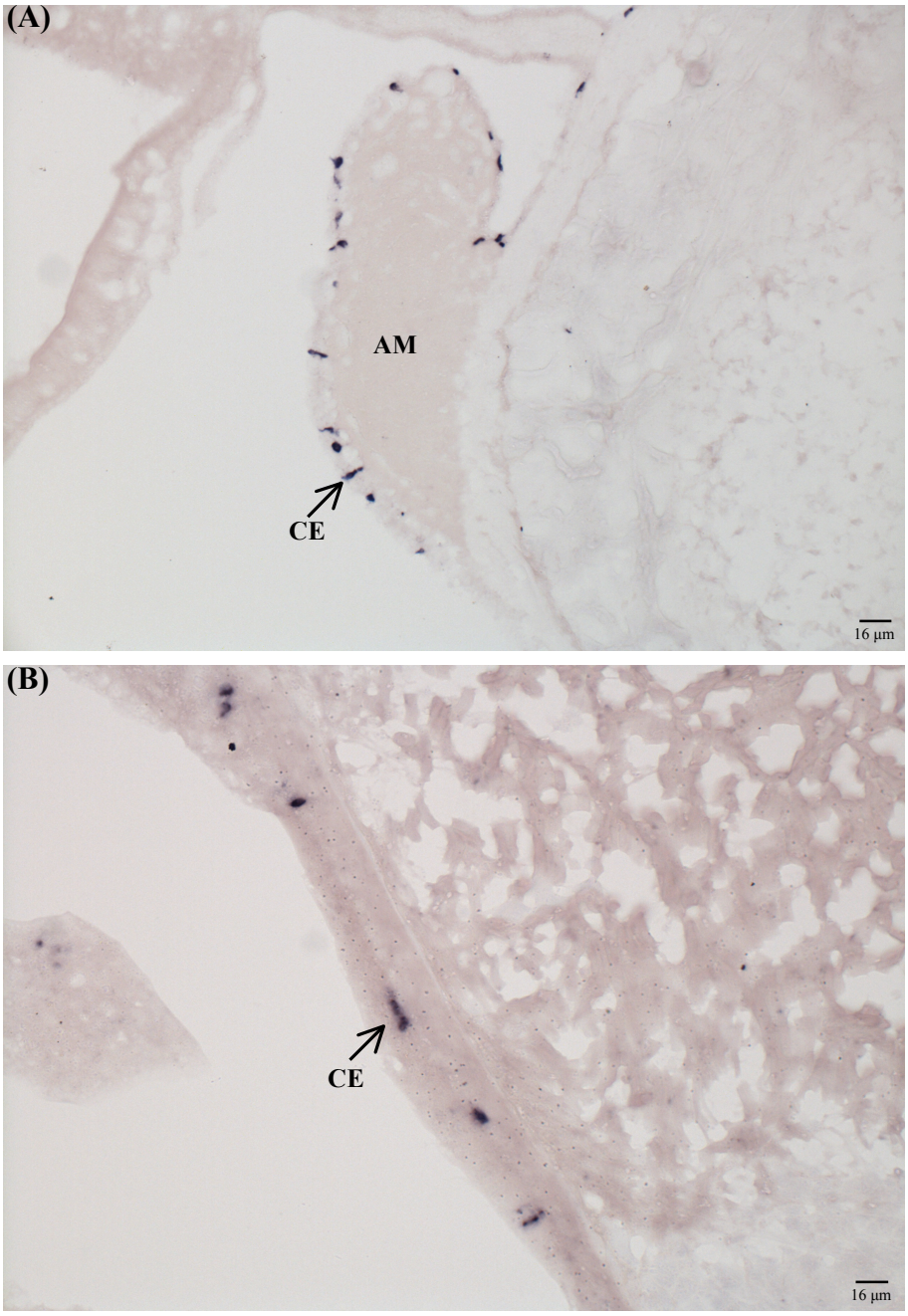
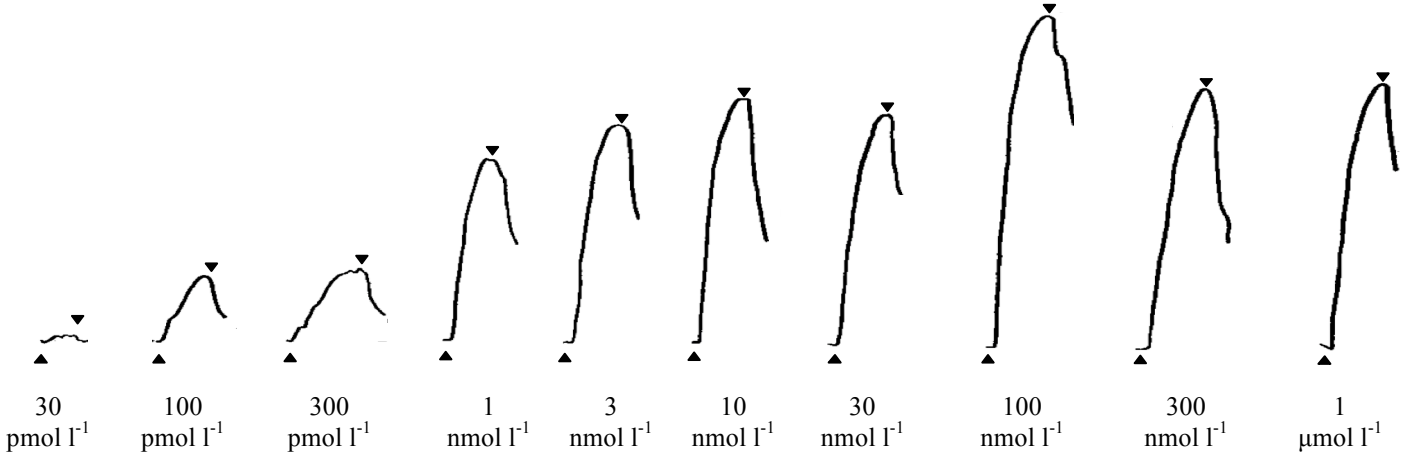


Fig. 4.9. NGFFYamide mRNA expression in the apical muscle and coelomic epithelium of *A. rubens*. (A) Transverse section through the central disk of *A. rubens* displaying the apical muscle. NGFFYamide mRNA expression was detected in the coelomic epithelial layer lining the apical muscle. **(B)** Transverse section through the central disk of *A. rubens* displaying the coelomic epithelium. NGFFYamide mRNA expression was detected in the coelomic epithelium. *AM = apical muscle*; *CE = coelomic epithelium.*

4.3.4. NGFFYamide causes cardiac stomach contraction *in vitro*

In order to determine whether synthetic NGFFYamide (NGFFY-NH_2) has an effect on cardiac stomach activity, an *in vitro* assay previously used to show that the SALMFamides S1 and S2 cause cardiac stomach relaxation was implemented (**Fig. 3.2**) (Elphick and Melarange, 1998; Elphick et al., 1995; Melarange et al., 1999). In the *in vitro* assay, the addition of NGFFYamide to cardiac stomach preparations revealed that it caused dose-dependent contraction at concentrations ranging from 30 pmol l^{-1} to 1 μmol l^{-1} (**Fig. 4.10**). A total of eight adult specimens were analysed, with the maximal efficacy of NGFFYamide observed at 100 nmol l^{-1} (**Fig. 4.10**). It should also be noted that synthetic NGFFFamide (NGFFF-NH_2) also caused dose-dependent contraction at concentrations ranging from 1 nmol l^{-1} to 1 μmol l^{-1}, with the maximal efficacy of NGFFFamide observed at 100 nmol l^{-1} (**Fig. 4.10B**). Comparison of the effects of NGFFYamide and NGFFFamide on cardiac stomach contraction using a random intercept linear mixed effects model (Bates and Sarkar, 2007) revealed a significant difference between both peptides on cardiac stomach contraction, irrespective of the concentration used ($P \leq 0.001$).

(A)


30 pmol l-1
100 pmol l-1
300 pmol l-1
1 nmol l-1
3 nmol l-1
10 nmol l-1
30 nmol l-1
100 nmol l-1
300 nmol l-1
1 µmol l-1

(B)

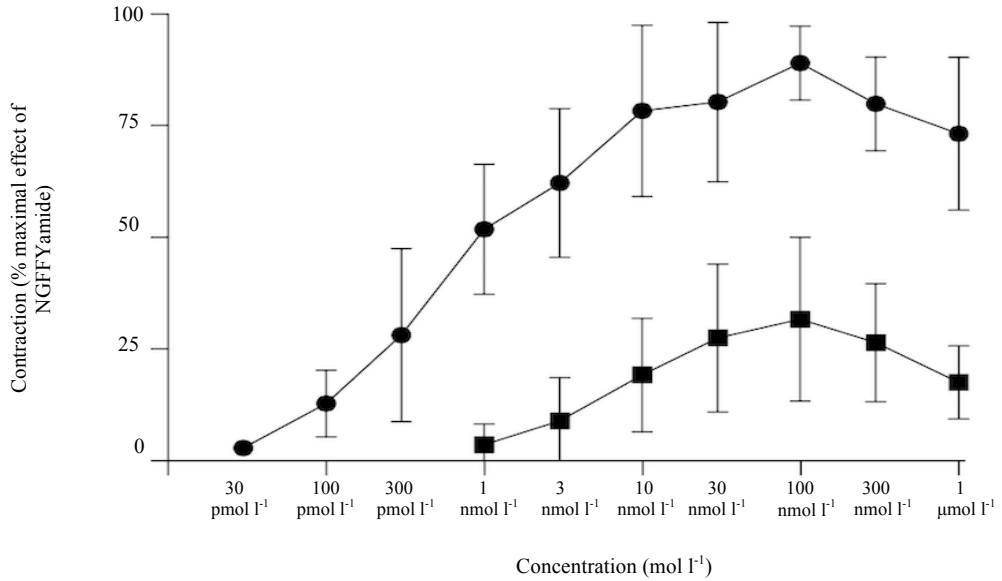
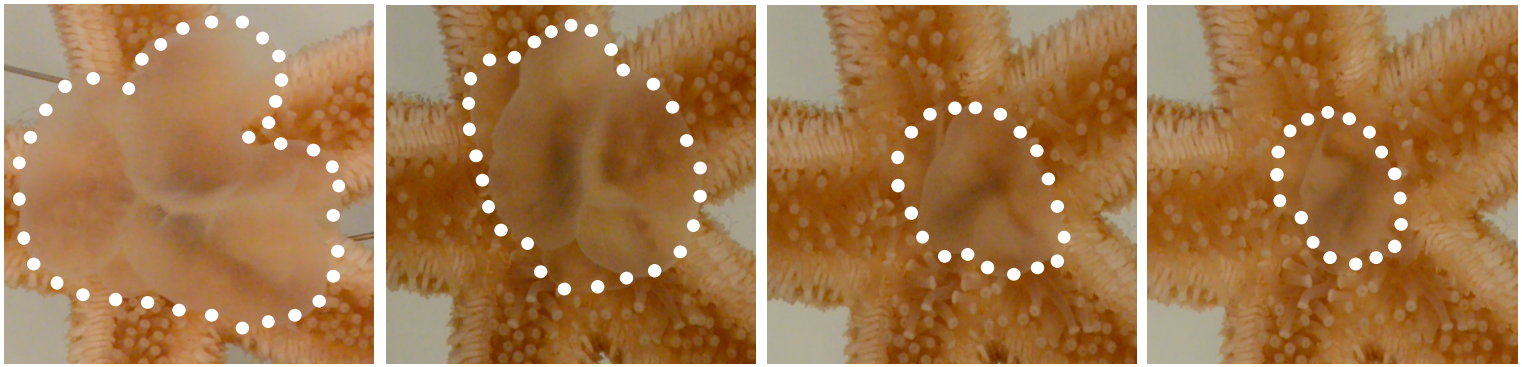

Contraction (% maximal effect of NGFFYamide)
100
75
50
25
0
30 pmol l-1
100 pmol l-1
300 pmol l-1
1 nmol l-1
3 nmol l-1
10 nmol l-1
30 nmol l-1
100 nmol l-1
300 nmol l-1
1 µmol l-1
Concentration (mol l-1)

Fig. 4.10. NGFFYamide is a potent stimulator cardiac stomach contraction in *A. rubens*. **(A)** Representative recordings from a single cardiac stomach preparation showing the dose-dependent effect of synthetic NGFFYamide (30 pmol l^{-1} to 1 μmol l^{-1}). The application of NGFFYamide (upward-pointing arrowheads) causes cardiac stomach contraction in a typical time-frame of 1-2 minutes, whilst the effect of NGFFYamide is reversed by washing (downward-pointing arrowheads). **(B)** Graph showing the dose-dependent effects of synthetic NGFFYamide (circles) and NGFFFamide (squares). The effects of both peptides are normalised to the maximal effect observed with NGFFYamide in each experiment, with mean values (± s.e.m.) from eight experiments shown.

4.3.5. NGFFYamide causes cardiac stomach retraction *in vivo*

In order to determine whether the effect of synthetic NGFFYamide (NGFFY-NH_2) on cardiac stomach contraction *in vitro* (**Fig. 4.10**) correlates with an effect on cardiac stomach contraction *in vivo*, cardiac stomach eversion was induced by placing starfish in a tank containing 2 % $MgCl_2$ (Mayer, 1909). In the *in vivo* assay, injection of NGFFYamide (10 µl of 100 nmol l^{-1}) into the perivisceral coelom triggered retraction of the cardiac stomach (**Fig. 4.11**). Informed by this initial observation, injection experiments were performed in a total of ten specimens in order to determine whether the effect of NGFFYamide (10 µl of 100 nmol l^{-1}) on cardiac stomach retraction was repeatable. NGFFYamide triggered cardiac stomach retraction in 10/10 experiments with variability in the rate and extent of retraction. **Fig. 4.11B** shows data from ten experiments, with the mean area of cardiac stomach everted at 20-second intervals during a 220 second recording period following peptide injection at time 0 (T_0) expressed as a percentage of the area everted at T_0. Importantly, in control experiments in which starfish were first injected with autoclaved water, no eversion of the cardiac stomach was observed.

(A)


0 sec
60 sec
120 sec
180 sec

(B)

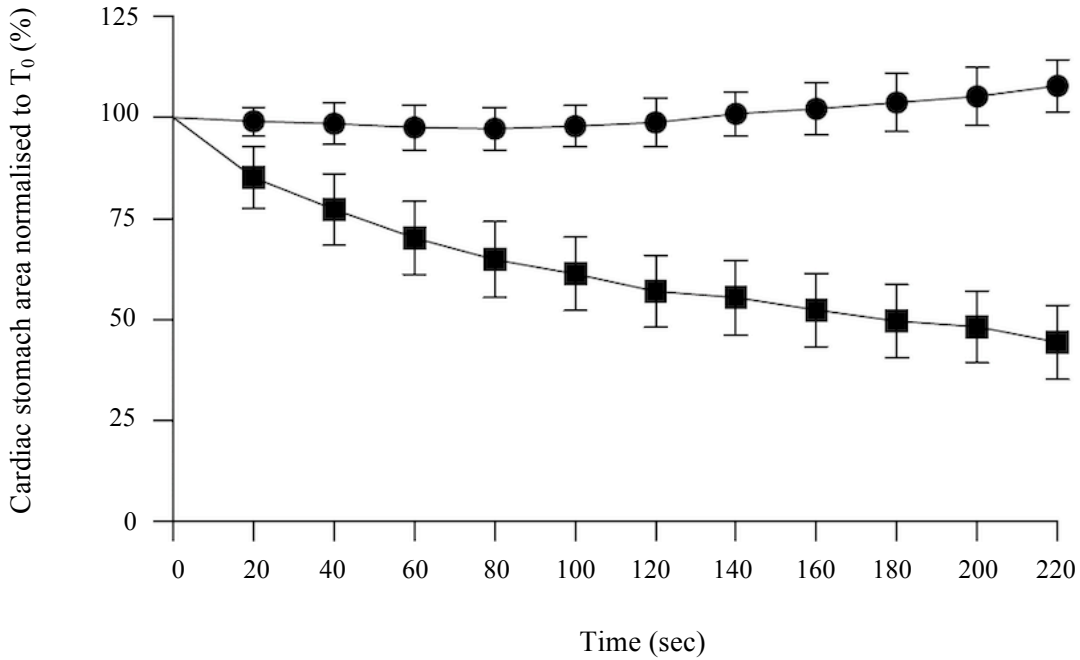

Cardiac stomach area normalised to T_0 (%)
125
100
75
50
25
0
0
20
40
60
80
100
120
140
160
180
200
220
Time (sec)

Fig. 4.11. NGFFYamide triggers cardiac stomach retraction in *A. rubens*. **(A)** Photographs from a representative experiment showing that injection of synthetic NGFFYamide (10 µl of 100 nmol l^{-1}) causes cardiac stomach retraction. At time 0 (T_0), the fully everted cardiac stomach can be seen. At 60 seconds, 120 seconds and 180 seconds after injection of NGFFYamide, cardiac stomach retraction is observed (outline of cardiac stomach marked by white dots). **(B)** Graph showing that injection of synthetic NGFFYamide (10 µl of 100 nmol l^{-1}) (squares) causes cardiac stomach retraction compared to injection of autoclaved water (10 µl) (circles). The area of cardiac stomach everted (in 2D) at each time point (0-220 seconds) is normalised to the area of the cardiac stomach everted at T_0, with means (± s.e.m.) from ten experiments shown.

4.4. Discussion

4.4.1. Discovery of NGFFYamide, an NG peptide in *A. rubens*

The NGFFYamide precursor in the starfish *A. rubens* is the first NG peptide precursor to be discovered in a starfish species. The NGFFYamide precursor contains an N-terminal signal peptide, two tandem copies of the sequence NGFFYG flanked by dibasic cleavage sites (KR) and a C-terminal neurophysin domain (**Fig. 4.2**). Importantly, mass spectrometry has confirmed that the C-terminal glycine (G) residue of the putative NG peptide is converted to an amide group (NGFFY-NH_2) (**Fig. 4.3**). In addition to the presence of a C-terminal amide group, comparison of the NGFFYamide precursor with other NG peptide precursors across the deuterostomian invertebrates reveals interesting similarities and differences. NGFFYamide, as with all known members of the NG peptide family, has a conserved asparagine-glycine (NG) motif. However, the position of the motif is variable. In NGFFYamide and other putative NG peptides in the echinoderms and hemichordates, the NG motif is found at the N-terminus of the putative peptide (e.g. NGFFYamide). However, in the cephalochordate *B. floridae*, the NG motif is not found at the N-terminus of the putative peptide (SFRNGVamide). The number of copies of putative NG peptides is also variable. For example, there are six copies of putative NG peptides in the hemichordate *S. kowalevskii*, five copies in the sea cucumber *A. japonicus* and two copies in the sea urchin *S. purpuratus* and the starfish *A. rubens*. Interestingly, comparison of spacer peptide sequences following copies of NGFWNamide and NGFYNamide in the NGFWNamide/NGFYNamide precursor in the hemichordate *S. kowalevskii* reveals sequence similarity indicative of an intragenic duplication (Elphick, 2010). However, whether the presence of multiple copies of putative NG peptides derived from the same precursor is

functionally significant remains open to investigation. Finally, the presence of a C-terminal neurophysin domain hitherto thought to be unique to VP/OT-type neuropeptide precursors appears to be a conserved feature of NG peptide precursors. Interestingly, however, the neurophysin domain is not present in the NGIWYamide precursor in the sea cucumber *A. japonicus*, indicating that it may have been lost in the holothurian lineage (Elphick, 2012).

4.4.2. NGFFYamide: a regulator of cardiac stomach contraction

Investigation into the *in vitro* pharmacological effects of NGFFYamide reveals that it causes dose-dependent relaxation of cardiac stomach preparations at concentrations ranging from 30 pmol l^{-1} to 1 μmol l^{-1}, with a maximal efficacy at 100 nmol l^{-1} (**Fig. 4.10**). Investigation into the *in vivo* effects of NGFFYamide revealed that it triggers cardiac stomach retraction in the starfish *A. rubens* (**Fig. 4.11**). Despite variability in the rate and extent of retraction, NGFFYamide triggered cardiac stomach retraction in 10/10 experiments. In combination, these data suggest that NGFFYamide may mediate neural control of cardiac stomach retraction during the process of feeding in starfish. Interestingly, NGFFYamide is the first neuropeptidergic mediator of cardiac stomach retraction to be discovered in the starfish *A. rubens*. However, the neurotransmitter ACh has previously been shown to cause *in vitro* biphasic contraction of the cardiac stomach in *A. rubens*, which suggests it may also cause cardiac stomach contraction *in vivo* (Elphick et al., 1995).

However, it is of interest to compare the effects of NGFFYamide to other neuropeptidergic mediators of the feeding process both *in vitro* and *in vivo*. *In vitro*, the maximal effect of NGFFYamide was observed at 100 nmol l^{-1} compared to the maximal effects of asterotocin at 100 nmol l^{-1} (**Fig. 3.16**) and the SALMFamides S1 and S2 at 10 μmol l^{-1} (Melarange et al., 1999). *In vitro*, NGFFYamide caused

cardiac stomach contraction at 30 pmol l^{-1} compared to asterotocin causing cardiac stomach relaxation at 30 pmol l^{-1} (**Fig. 3.16**) and the SALMFamides S1 at 10 nmol l^{-1} and S2 at 1 nmol l^{-1} respectively (Melarange et al., 1999). *In vivo*, NGFFYamide triggered cardiac stomach retraction upon injection of 10 µl of 100 nmol l^{-1} peptide compared to the injection of 10 µl of 10 µmol l^{-1} asterotocin (**Fig. 3.14**) and 100 µl of 1 mmol l^{-1} S1 and S2 in triggering cardiac stomach eversion (Melarange et al., 1999). The differences in the concentration of peptide required to elicit a response on myoactivity could be in part due to the relative nature of access of the peptide to muscles in the cardiac stomach. In this regard, NGFFYamide is injected when the cardiac stomach is already fully everted whilst asterotocin is injected when access to the cardiac stomach muscle is limited by the compact nature of the cardiac stomach in the central disk prior to eversion.

In combination with the effects of NGFFYamide *in vitro* and *in vivo*, the expression profile of NGFFYamide mRNA provides further evidence for the role of NGFFYamide in mediating neural control of cardiac stomach retraction. Importantly, NGFFYamide mRNA is expressed in the cardiac stomach of the starfish *A. rubens*, with cell bodies concentrated in the basiepithelial nerve plexus, with fewer cell bodies associated with the and the mucosal layer and visceral muscle layer (**Fig. 4.7**). This expression profile suggests that NGFFYamide may mediate cardiac stomach retraction by two potential mechanisms. Firstly, NGFFYamide may be released from processes of NGFFYamide-expressing cells located in the basiepithelial nerve plexus and diffuses across a thin connective tissue layer to reach the visceral muscle layer (**Fig. 3.10**). Alternatively, processes of NGFFYamide-expressing cells may project across the connective tissue layer to reach the visceral muscle layer directly. In this regard, developing antibodies to NGFFYamide and performing immunocytochemistry (ICC) will be crucial in providing insights into the functional

mechanism of the peptide.

Interestingly, it should be noted that cell bodies appear to be more densely concentrated in the aboral region as opposed to the oral region of the cardiac stomach, which is densely packed with granular secretory cells (**Fig. 4.7**). It should also be noted that NGFFYamide mRNA was also detected in the pyloric stomach and pyloric ducts, with cell bodies sparsely located in the basiepithelial nerve plexus and mucosal layer but absent from the visceral muscle layer (**Fig. 4.8**). Similarly to the mRNA expression profile for asterotocin (see **chapter 3.3.3**), the mRNA expression profile of NGFFYamide throughout the digestive system reinforces the role of NGFFYamide in regulating cardiac stomach retraction as opposed to digestive processes.

It will be important to first determine where NGFFYamide and, upon pharmacological characterisation, the candidate NGFFYamide receptor(s) are expressed in relation to one another in the context of the starfish body plan to further explore the role of NGFFYamide in cardiac stomach retraction. In addition to utilising mRNA ISH, it will also be important to develop antibodies to NGFFYamide and the candidate NGFFYamide receptor to perform ICC, which will help to distinguish the specific cell types that NGFFYamide is expressed in with regards to the mucosal layer of the gut, which includes ciliated epithelial cells (which propel food particles from the cardiac stomach through to the pyloric stomach) in addition to mucous goblet cells and granular secretory cells (Moore and Thorndyke, 1993). The expression of NGFFYamide mRNA in the cardiac stomach (and absence from retractor strands) suggests that NGFFYamide may exert its effects on the cardiac stomach either directly or indirectly (via cell bodies in the basiepithelial nerve plexus) by binding to its candidate receptor(s), which may lead to changes in secondary messenger levels or ion channel conductance (Elphick and Melarange,

2001). However, there remains the possibility that NGFFYamide may act indirectly to innervate the visceral muscle layer. It will therefore be crucial to determine the signalling cascades that are involved in eliciting the cardiac stomach retraction response in order to provide insights into the peptide's mode of action. In this regard, determining the receptor type that the NG peptide family activate and the signal transduction pathways that the peptides couple to will be key (see **chapter 5**).

4.4.3. The NG peptides: a conserved role in muscle myoactivity?

The discovery of NGFFYamide has opened up the possibility to investigate the range of functions that the peptide may cover in the starfish *A. rubens*. Previous functional studies in the echinoderms, although limited, have revealed that the NG peptides may have a conserved role in regulating a range of physiological functions associated with contractile activity.

In the sea urchin species *E. esculentus*, NGFFFamide has been shown to cause contraction of *in vitro* tube foot and oesophagus preparations (Elphick and Rowe, 2009). However, NGFFFamide is a more potent contractant of oesophagus preparations as opposed to tube feet preparations (Elphick and Rowe, 2009). In the sea cucumber *A. japonicus*, NGIWYamide has been shown to have a variety of contractile effects. NGIWYamide has been shown to cause contraction of RLM and intestinal muscle preparations (Iwakoshi et al., 1995; Ohtani et al., 1999), tentacle preparations (Inoue et al., 1999) and control the stiffness of connective tissue in the body wall dermis (Birenheide et al., 1998). More recently, it has also been shown that NGIWYamide induces oocyte maturation and gamete spawning in the sea cucumber *A. japonicus* (Kato et al., 2009). Combined with functional data in the sea urchin species *E. esculentus* and the starfish *A. rubens*, the NG peptides appear to regulate a range of physiological functions associated with contractile activity.

Interestingly, the contractile effect of NGFFFamide on oesophagus preparations from the sea urchin species *E. esculentus* (Elphick and Rowe, 2009) mirrors the contractile role of NGFFYamide on cardiac stomach preparations from *A. rubens*, which is suggestive of a role for the NG peptides in the feeding process in the echinoderms. However, it will important to determine whether NGFFYamide acts as a general muscle contractant on a range of preparations from *A. rubens* before drawing any hypotheses on a conserved role for the NG peptides.

In support of a general myoactive role for NGFFYamide, NGFFYamide mRNA was detected in the nerve plexus in the adhesive region of the tube feet, in the basal nerve ring and was also present in marginal nerve, which innervates the outer row of tube feet in the starfish *A. rubens* (**Fig. 4.6**). These findings are consistent in part with NGIWYamide immunoreactivity in the starfish species *A. pectifinera*, in which expression was also observed in the basal nerve ring, marginal nerve and the nerve plexus of the adhesive region of the tube foot (Saha et al., 2006). The expression of NGFFYamide mRNA was observed in cell bodies in the coelomic epithelial layer lining the apical muscle and the coelomic epithelium separating the body wall from the coelomic cavity (**Fig. 4.9**). Interestingly, initial functional studies on *in vitro* tube feet and apical muscle preparations reveal that NGFFYamide causes dose-dependent relaxation of both preparations (Dane and Elphick, unpublished data). These data are the first in an echinoderm species that show that an NG peptide can cause relaxing effects as opposed to solely contractile-type effects. Once more, it will be crucial to determine the receptor(s) that the NG peptide family activate to determine the signal transduction pathways that lead to both the contractile and relaxing-type effects of NGFFYamide.

4.4.4. The NG peptides: neurophysin-associated neuropeptides

Neurophysins are highly conserved pro-hormone derived polypeptides required for sorting the neurohypophyseal hormones VP and OT into the RSP (Cool et al., 2008; de Bree, 2000; de Bree and Burbach, 1998). The importance of the neurophysin domain can be highlighted through the disease FNDI, which is caused by mutations in the neurophysin domain of human VP, leading to an impairment of VP biosynthesis (Babey et al., 2011).

The presence of a neurophysin domain was previously thought to be unique to VP/OT-type neuropeptide precursors. However, the discovery that NG peptide precursors have a C-terminal neurophysin domain has opened up the possibility to investigate the function of the neurophysin domain in NG peptide biosynthesis. However, there remains the possibility that the neurophysin domain is not required for NG peptide biosynthesis, which may explain why the neurophysin domain has been lost in the NGIWYamide precursor in the sea cucumber *A. japonicus* (Elphick, 2012).

5. Discovery of the NGFFFamide receptor in the sea urchin *S. purpuratus* unites a bilaterian neuropeptide family

5.1. Introduction

Neuropeptides can be found in highly conserved families and can therefore be readily identified in evolutionary distant animals within the animal kingdom. However, such high sequence conservation is relatively uncommon and therefore establishing evolutionary relationships between neuropeptides in evolutionarily distant phyla has historically proved difficult. Recently, with advances in comparative genomics/transcriptomics, a number of neuropeptide signalling systems have been shown to be of an ancestral bilaterian origin (Jekely, 2013; Mirabeau and Joly, 2013). In many cases, these associations have only been made possible by investigating neuropeptide signalling systems in a range of phyla across the animal kingdom. Importantly, a vast number of these evolutionary insights have been made possible on the basis on receptor orthology.

5.1.1. Neuropeptide-S (NPS)

The sequencing of the human genome led to the identification of a vast number of genes encoding G-protein coupled receptors (GPCRs) with unknown endogenous ligands – these GPCRs were coined “orphan receptors”. One of the first receptors to be deorphanised through identification of its endogenous ligand was GPR154. The receptor was identified as being closely related to vasopressin/oxytocin (VP/OT)-type receptors and was therefore referred to as VP receptor-related receptor 1 (VRR1), mapping to chromosome 7p14-15 and displaying selective expression in the hypothalamus and retina (Gupte et al., 2004). However, through subsequent

pharmacological characterisation, it was found that VP/OT-type peptides do not act as ligands at VRR1. Instead, it was found that the endogenous ligand for VRR1 is a 20-residue neuropeptide (SFRNGVGTGMKKTSFQRAKS) termed neuropeptide-S (NPS) on account of its N-terminal serine (S) residue (Xu et al., 2004). Interestingly, despite activating a VP/OT-type receptor, NPS does not share any sequence similarity with VP/OT-type peptides.

NPS was first discovered in mammals and has subsequently been identified as belonging to a highly conserved family of neuropeptides in tetrapod vertebrates (Reinscheid, 2007). NPS modulates arousal and anxiolytic-like behaviour in rodents (Xu et al., 2004), with polymorphisms in the NPS receptor (NPSR) also associated with panic disorders in humans (Domschke et al., 2011; Okamura et al., 2007). Taken together, these findings indicate that NPS is an endogenous regulator of anxiety in tetrapod vertebrates.

5.1.2. Crustacean cardioactive peptide (CCAP)

Investigation into the phylogenetic distribution of NPSR-like proteins reveals that orthologues of the NPSR are not only found in tetrapod vertebrates but also are present in a number of invertebrate species including annelids, molluscs, insects and crustaceans (Jekely, 2013; Mirabeau and Joly, 2013). The endogenous ligand found to activate an NPSR-like protein in *Drosophila* was found to be a 9-residue amidated neuropeptide (PFCNAFTGC-NH_2) termed crustacean cardioactive peptide (CCAP) (Cazzamali et al., 2003; Park et al., 2002). Interestingly, despite activating an NPSR-like protein, CCAP does not share any sequence similarity with NPS-type peptides.

CCAP was first discovered in the crustacean *Carcinus maenas* (Stangier et al., 1987) and has subsequently been identified as belonging to a highly conserved family of neuropeptides in protostomian invertebrates (Jekely, 2013; Mirabeau and

Joly, 2013). CCAP was initially discovered as controlling cardioacceleration in the crustacean *C. maenas* (Stangier et al., 1987) and has subsequently found to have a key role in neural control of ecdysis behaviour in crustaceans (Chung and Webster, 2004; Phlippen et al., 2000) and insects (Arakane et al., 2008; Belmont et al., 2006; Ewer et al., 1998; Gammie and Truman, 1997; Kim et al., 2006; Park et al., 2003).

It has been proposed that CCAP shares superficial sequence similarity with VP/OT-type peptides in the form of two cysteine (C) residues forming a disulphide bridge identified as being crucial for bioactivity (Arakane et al., 2008; Stangier et al., 1987). However, CCAP precursor proteins do not have a C-terminal neurophysin domain characteristic of VP/OT-type precursor proteins (Loi et al., 2001). Interestingly, CCAP receptors are paralogous to VP/OT-type receptors, suggesting that CCAP and VP/OT-type peptides may have evolved from a common ancestral neuropeptide system (Mirabeau and Joly, 2013). However, despite paralogous receptors, what remains unclear is the relationship between NPS and VP/OT-type peptides and CCAP, with NPS sharing no apparent sequence or structural similarity with either VP/OT-type peptides or CCAP.

5.1.3. The NG peptides

Recent advances in comparative genomics/transcriptomics have transformed our understanding of neuropeptide diversity across the animal kingdom. A strategy to bridge the gap of understanding between neuropeptide signalling systems in vertebrates and classic model protostomian invertebrates (e.g. *Drosophila* and *Caenorhabditis elegans*) is to investigate genomic/transcriptomic data from the deuterostomian invertebrates (e.g. cephalochordates, hemichordates and echinoderms). Investigation into the phylogenetic distribution of NPSR-like proteins in deuterostomian invertebrate genome data reveals that orthologues of NPS/CCAP-

type receptors are also present in the cephalochordates, hemichordates and echinoderms (Mirabeau and Joly, 2013; Pitti and Manoj, 2012; Valsalan and Manoj, 2014).

The NG peptides are a novel family of neuropeptides characterised by an asparagine-glycine (NG) motif that have recently been identified in deuterostomian invertebrates (Elphick, 2010). The first NG peptide precursor to be discovered was the NGFFFamide precursor in the sea urchin *S. purpuratus*, which has two tandem copies of the neuropeptide NGFFFamide (Elphick and Rowe, 2009). Subsequently, NG peptide precursors have been identified in a number of other echinoderm species (see **chapter 4**), the hemichordate *Saccoglossus kowalevskii* and the cephalochordate *Branchiostoma floridae* (Elphick, 2010). Interestingly, the NG peptide precursor in *B. floridae* has two copies of the putative NG peptide SFRNGVamide that shares sequence identity with the N-terminal region of NPS (SFRNGVGTGMKKTSFQRAKS), which is crucial for its bioactivity (Roth et al., 2006); this is indicative of a common evolutionary origin (Elphick, 2010). Interestingly, NG peptide precursors have a C-terminal neurophysin domain, which was previously believed to be uniquely associated with VP/OT-type precursors (de Bree, 2000; de Bree and Burbach, 1998). Thus, with NPS and VP/OT-type peptides acting as ligands for orthologous receptors, the discovery of the neurophysin-associated NG peptides has provided a link between NPS and VP/OT-type peptides.

5.1.4. NPS-type peptides, NG peptides and CCAP-type peptides: a bilaterian neuropeptide family?

Taken together, it appears that structural features characteristic of the VP/OT-type neuropeptide signalling system have been differentially retained in the protostomian and deuterostomian invertebrates. In protostomian invertebrates, it is the presence of a disulphide bridge between two cysteine (C) residues in CCAP that is characteristic of VP/OT-type peptides. In deuterostomian invertebrates, despite sharing no sequence similarity with VP/OT-type peptides, it is the presence of a C-terminal neurophysin domain in the NG peptide precursors that is characteristic of VP/OT-type precursors. Finally, despite NPS precursors not containing a C-terminal neurophysin domain, it is sequence similarity shared between the N-terminal region of NPS and SFRNGVamide in the cephalochordate *B. floridae* that is indicative of a common evolutionary origin (Elphick, 2010). Therefore, it can be deduced that the neurophysin domain was lost in an NG peptide-type precursor that gave rise to NPS precursors in the tetrapod vertebrates.

Collectively, these observations indicate that duplication of a VP/OT-type neuropeptide precursor gene may have occurred in the common ancestor of the Bilateria. Here, one copy was retained giving rise to the VP/OT-type neuropeptide precursors found throughout extant bilaterians. The second copy diverged and took different courses during evolution in the protostomian and deuterostomian invertebrates. Moreover, the remaining gene copy diverged throughout bilaterian evolution giving rise to the CCAP system in protostomes and the NG peptide and NPS systems in deuterostomes respectively.

Importantly, evidence in support of this model of neuropeptide evolution has been obtained through analysis of gene contiguity in the genomes of the deuterostomian invertebrates *S. purpuratus* and *B. floridae*. In the echinoderm *S.*

purpuratus, a gene encoding a VP/OT-type receptor (SPU_021290) is located adjacent to a gene encoding a candidate NPS/CCAP-type receptor (SPU_021291) (**Fig. 5.1**). In the cephalochordate *B. floridae,* a gene encoding a VP/OT-type precursor (Brafl1-84802) is located adjacent to a gene encoding the SFRNGVamide precursor (Brafl1-84803) (Mirabeau and Joly, 2013). However, to date, such gene contiguities have not been identified in other species, potentially reflecting gene rearrangement within these species.

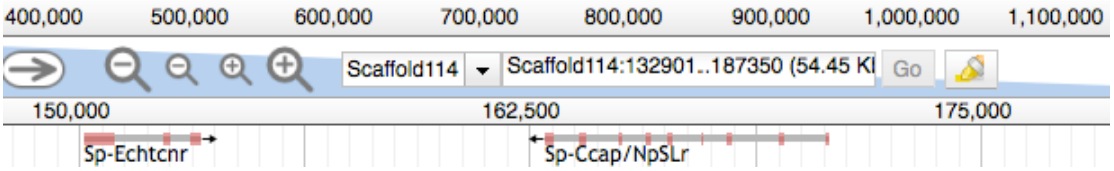


Fig. 5.1. Gene contiguity in the sea urchin *S. purpuratus*. Diagram showing that a VP/OT-type ("echinotocin") receptor gene (Sp-Echtcnr) (SPU_021290) is located adjacent to a NPS/CCAP-type receptor gene (Sp-Ccap/NpSLr) (SPU_021291) on scaffold 114 of the *S. purpuratus* genome. Exons are represented by red rectangles and introns are represented by grey rectangles. The diagram was generated as an output from the *S. purpuratus* genome browser (http://www.echinobase.org/Echinobase/Spbase) utilising the JBrowse Genome Browser v.1.10.9 (http://jbrowse.org).

5.1.5. Aims and objectives

The primary aim of the work reported in this chapter is to clone, sequence and pharmacologically characterise the receptor for the NG peptide NGFFFamide in the sea urchin *S. purpuratus*. The candidate NGFFFamide receptor is orthologous to NPS-type receptors in the tetrapod vertebrates and CCAP-type receptors in the protostomian invertebrates. Thus, characterisation of the NGFFFamide receptor will unify a bilaterian neuropeptide family that includes NPS-type peptides in tetrapod vertebrates, NG peptides in deuterostomian invertebrates and CCAP-type peptides in protostomian invertebrates.

The secondary aim of this chapter is to identify NG peptide precursors and candidate NPS/CCAP-type receptors utilising recently available genome/transcriptome data from a number of echinoderm species. In combination, these data will provide an important foundation to pharmacologically characterise NG peptide receptors across the echinoderms.

5.2. Methods

5.2.1. Bioinformatic identification of an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*

A candidate neuropeptide receptor (SPU_021291) was identified through the sea urchin genome project (Burke et al., 2006; Sodergren et al., 2006) and has subsequently been identified as an NPS/CCAP-type receptor on account of sequence similarity with NPS/CCAP-type receptors in both the protostomian and deuterostomian invertebrates (Mirabeau and Joly, 2013; Valsalan and Manoj, 2014).

5.2.2. Cloning and sequencing of an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*

S. purpuratus cDNA (72-hour embryonic) generated using an iScript^{TM} cDNA Synthesis Kit (Bio-Rad) was kindly provided by Dr. Paola Oliveri (UCL). NPS/CCAP-type receptor cDNA containing the entire coding region was amplified by PCR using Phusion High-Fidelity PCR Master Mix (NEB) and the oligos: 5'-CCGACATAGAAAGTCATAG-3' and 5'-GTCAGATCAGATATAGCAGT-3'. The PCR product was gel-extracted and purified using the QIAquick Gel Extraction Kit (QIAgen) before being blunt-end cloned into a pBluescript SKII (+) vector (Agilent Technologies) and cut with the EcoRV-HF restriction endonuclease (NEB). Upon sequencing a clone from both the T3 and T7 sequencing primer sites for sequence confirmation (Eurofins), the full-length cDNA of the *S. purpuratus* NPS/CCAP-type receptor was sub-cloned into the eukaryotic expression vector pcDNA 3.1+ (Invitrogen) cut with BamHI and ApaI restriction endonucleases (NEB) using the oligos: 5'-cgggatccCACCATGGCGACACAAGTGA-3' and 5'-atcgggcccCTACATCGGACTAGTAGTC-3'. The partial Kozak sequence (CACC)

was incorporated immediately preceding the start codon to optimise initiation of translation. The clone was then sequenced from the T7 and pCR3.1-BGH-rev sequencing primer sites (Eurofins). A plasmid maxiprep was subsequently performed using the QIAgen Plasmid Maxi Kit (QIAgen).

5.2.3. Pharmacological characterisation of heterologously expressed candidate NGFFFamide receptor

Chinese hamster ovary (CHO) K1 (CHO-K1) cells stably overexpressing the mitochondrially targeted apo-aequorin and the promiscuous human $G\alpha_{16}$ protein were used as a heterologous expression system for an *in vitro* Ca^{2+} mobilisation assay as has been employed previously for identification of neuropeptide receptors in the nematode *C. elegans* (Janssen et al., 2008). CHO-K1 cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM)/F12-Ham (Sigma-Aldrich) supplemented with 10 % fetal bovine serum (FBS) (Sigma-Aldrich), 100 U/ml of penicillin/streptomycin (Invitrogen), 250 µg/ml zeocin (Invitrogen) and 2.5 µg/ml amphotericin B (Sigma-Aldrich), in a humidified atmosphere of 5 % carbon dioxide (CO_2) at 37 °C. CHO-K1 cells were transiently transfected with the pcDNA 3.1+ eukaryotic expression vector containing the NPS/CCAP-type receptor from the sea urchin *S. purpuratus* using Lipofectamine® LTX with Plus reagent™ (Invitrogen) according to the manufacturer's instructions. CHO-K1 cells for negative control experiments were transfected with an empty pcDNA 3.1+ eukaryotic expression vector (Invitrogen). Two days post-transfection, cells were challenged with the synthetic NG peptides NGFFFamide (NGFFF-NH_2) (*S. purpuratus*), NGFFYamide (NGFFY-NH_2) (*Asterias rubens*) and NGIWYamide (NGIWY-NH_2) (*A. japonicus*) within the concentration ranges 10^{-17} M to 2.5 x 10^{-4} M. Potential activation of the NPS/CCAP-type receptor at each peptide concentration was monitored by measuring

Ca^{2+} levels for 30 seconds using a Mithras LB 940 luminometer (Berthold Technologies). Ca^{2+} responses were normalised to the total Ca^{2+} response monitored after addition of Triton X-100 (0.1 %), and measured in triplicate in at least two independent experiments. Dose-response data were determined as relative (%) to the highest value (100 % activation). Half maximal effective concentrations (EC_{50} values) were calculated from dose-response curves, constructed with non-linear regression analysis using a sigmoidal dose-response equation on SigmaPlot v.12.0 (Systat Software Inc.).

5.2.4. Bioinformatic identification of novel NG peptide-type precursors and candidate NG peptide receptors in the echinoderms

To search for transcripts encoding NG peptide precursors and NPS/CCAP-type receptors that have not previously not identified in the echinoderms, the NGFFFamide precursor protein sequence (Elphick and Rowe, 2009) and the NPS/CCAP-type receptor protein sequence (Mirabeau and Joly, 2013; Valsalan and Manoj, 2014) from the sea urchin *S. purpuratus* (or the NGIWYamide precursor from the sea cucumber species *A. japonicus* for *Parastichopus parvimensis* due to loss of a C-terminal neurophysin domain) were submitted as queries in a tBLASTn search of a number of genome/transcriptome datasets: the publically available genome of the sea urchin species *Lytechinus variegatus* (http://www.echinobase.org/Echinobase/Blast/LvBlast/blast.php); transcriptomes of the sea cucumber species *A. japonicus* (Du et al., 2012) and *P. parvimensis* (http://www.echinobase.org/Echinobase/PBase); the transcriptome of the starfish species *A. rubens* (see **chapter 2**) and the publically available genome of *P. miniata* (http://www.echinobase.org/Echinobase/pm); transcriptomes of the brittle star species *Ophionotus victoriae* (Elphick et al., 2015) and *Amphiura filiformis* (access

to this transcriptome dataset was provided by Dr. Paola Oliveri, UCL); transcriptomes of the feather star species *Antedon mediterranea* (Elphick et al., 2015) and *Florometra serratissima* (access to this transcriptome dataset was provided by Dr. Christopher Lowe, Stanford University).

In the NG peptide precursor queries, the top-hits (expected (e)-value) for each species were as follows: *L. variegatus* (scaffold 1378; 96196 base pairs; e-value: $1e^{-53}$) encoding a 265-residue protein; *P. parvimensis* (contig 177728; 1825 base pairs; e-value: $2e^{-165}$) encoding a 239-residue protein; *P. miniata* (contig 387981; 3147 base pairs; e-value: $4e^{-16}$) encoding a 239-residue protein; *O. victoriae* (contig 2050454; 1597 base pairs; e-value: $6e^{-21}$) encoding a 248-residue protein; *A. filiformis* (contig 3155347; 610 base pairs; e-value: $8e^{-31}$) encoding a partial 182-residue protein. These proteins were analysed for N-terminal signal peptides using SignalP 3.0 (Bendtsen et al., 2004), putative NG peptides flanked by dibasic cleavage sites and a C-terminal neurophysin domain, which is characterised by the presence of fourteen cysteine (C) residues (Elphick, 2010; Elphick and Rowe, 2009). Subsequently, protein sequences were submitted as tBLASTn searches of The National Centre for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine sequence similarity to known neuropeptide or peptide hormone precursors.

In NPS/CCAP-type receptor queries, the top-hits (e-value) for each species were as follows: *L. variegatus* (scaffold 255; 274595 base pairs; e-value: $4e^{-33}$) encoding a 440-residue protein; *A. japonicus* (isotig14190; 1646 base pairs; e-value: $2e^{-113}$) encoding a 423-residue protein; *P. parvimensis* (contig 183467; 2956 base pairs; e-value: $8e^{-109}$) encoding a 423-residue protein; *A. rubens* (contig 1116923; 1994 base pairs; e-value: $2e^{-118}$) encoding a 449-residue protein; *P. miniata* (scaffold 44; 185723 base pairs; e-value: $3e^{-22}$) encoding a 439-residue protein; *A.*

mediterranea (contig 1796804; 1746 base pairs; e-value: $7e^{-98}$) encoding a 394-residue protein; *F. serratissima* (contig 26548l; 2182 base pairs; e-value: $1e^{-95}$) encoding a 332-residue protein. These proteins were analysed for seven transmembrane domains characteristic of GPCRs using GPCRHMM (http://gpcrhmm.sbc.su.se). Subsequently, protein sequences were submitted as tBLASTn searches of The National Centre for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine sequence similarity to known neuropeptide or peptide hormone receptors.

5.3. Results

5.3.1. Cloning and sequencing of an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*

An NPS/CCAP-type receptor identified in the genome of the sea urchin *S. purpuratus* has been reported previously (Burke et al., 2006; Mirabeau and Joly, 2013; Valsalan and Manoj, 2014) and its sequence has subsequently been confirmed through transcriptome sequencing (Tu et al., 2012). However, in order to test the hypothesis that NGFFFamide is the endogenous ligand for an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*, the 444-residue sequence was cloned and sequenced (**Fig. 5.2**). The cloned and sequenced receptor contained a number of polymorphisms at the nucleotide level, but was identical at the protein level to previously reported genome (Burke et al., 2006) and transcriptome (Tu et al., 2012) data.

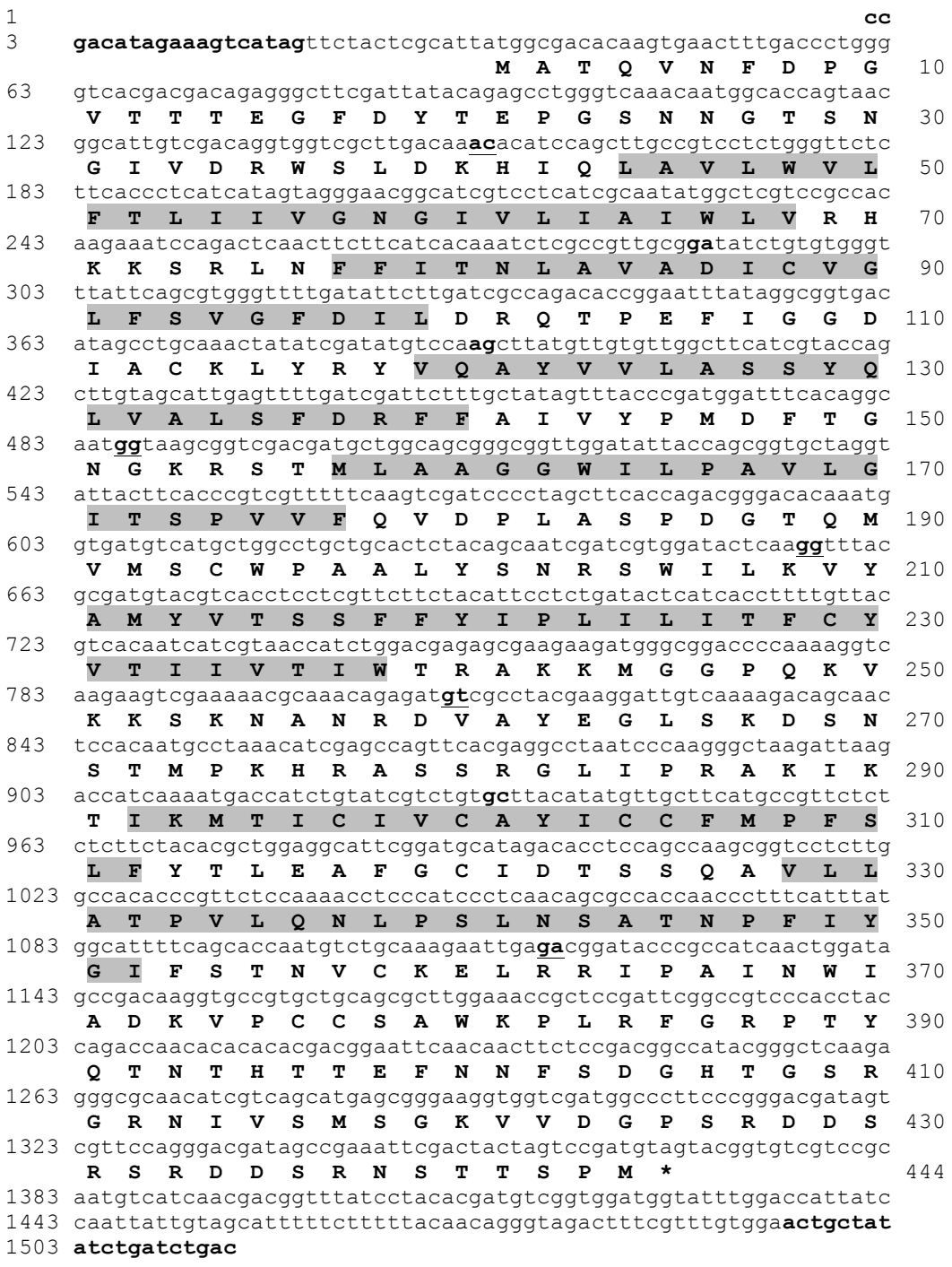


Fig. 5.2. ***S. purpuratus*** **NPS/CCAP-type receptor.** The nucleotide sequence (lowercase, 1514 bases) encoding the receptor protein (uppercase, 444 amino acid residues) confirmed by cloning and sequencing is shown, with primer sites shown by highlighting bases in bold. The positions of introns in the gene are shown by highlighting the pairs of bases (bold and underline) in the sequence that are interrupted by an intron. The asterisk shows the position of the stop codon. Predicted transmembrane domains are highlighted in grey.

5.3.2. NGFFFamide is a potent ligand for an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*

In order to test the hypothesis that the endogenous ligand for an NPS/CCAP-type receptor in the sea urchin *S. purpuratus* is the NG peptide NGFFFamide, an approach that has been successfully used to deorphanise a number of GPCRs in the nematode *C. elegans* was implemented (Janssen et al., 2008). The cloned and sequenced NPS/CCAP-type receptor was expressed in a CHO-K1 cell line stably overexpressing the human Gα16 protein (directing intracellular signalling via the phospholipase-C (PLC) and inositol triphosphate (IP_3) pathway of Ca^{2+} release regardless of endogenous G-protein coupling) and the mitochondrially-targeted apo-aequorin (to monitor the subsequent intracellular Ca^{2+} elevation). The NPS/CCAP-type receptor was activated in a dose-dependent manner by synthetic NGFFFamide (NGFFF-NH_2) with an EC_{50} of 0.4 nM (**Fig. 5.3**). In order to characterise structure-activity relationships of the NPS/CCAP-type receptor, synthetic NGFFYamide (NGFFY-NH_2) (*A. rubens*; see **chapter 4**) and synthetic NGIWYamide (NGIWY-NH_2) (*A. japonicas*; Iwakoshi et al., 1995; Elphick, 2012) were also investigated as potential ligands for the receptor. The NPS/CCAP-type receptor was activated by synthetic NGFFYamide and NGIWYamide with an EC_{50} of 5.7 nM and 367 nM respectively (**Fig. 5.3**). It is noteworthy that there is over a ten-fold difference in potency between NGFFFamide and NGFFYamide despite differing by only one amino acid and more specifically a single hydroxyl group on the C-terminal tyrosine (Y) residue of NGFFYamide. It is also noteworthy that the additional fold difference in potency between NGFFYamide and NGIWYamide is likely due to differences in the side chain structures of the isoleucine-tryptophan (IW) dipeptide as opposed to the phenylalanine-phenylalanine (FF) dipeptide present in NGFFFamide and NGFFYamide.

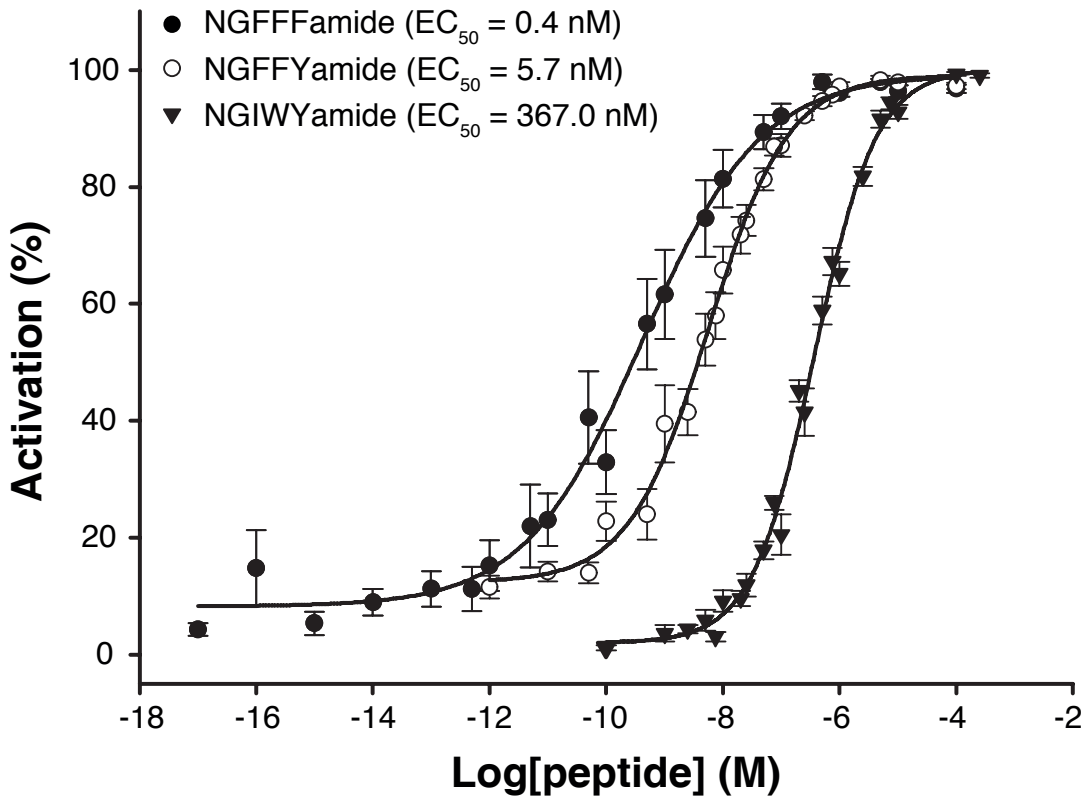


Fig. 5.3. The *S. purpuratus* NPS/CCAP-type receptor is activated by the NG peptides NGFFFamide, NGFFYamide and NGIWYamide. CHO-K1 cells expressing the *S. purpuratus* NPS/CCAP-type receptor were challenged with the synthetic NG peptides NGFFFamide (NGFFF-NH_2) (*S. purpuratus*), NGFFYamide (NGFFY-NH_2) (*A. rubens*) and NGIWYamide (NGIWY-NH_2) (*A. japonicus*) and Ca^{2+} response measured. Each point (± s.e.m) represents mean values from at least two independent experiments performed at least in triplicate. Dose-response data are shown as relative (%) to the highest value (100 % activation) after normalisation to the maximum Ca^{2+} response. The log EC_{50} values (± s.e.m) are NGFFFamide - 9.38 ± 0.09; NGFFYamide: - 8.25 ± 0.04; NGIWYamide: - 6.44 ± 0.04.

5.3.3. Sequences of novel NG peptide precursors and candidate receptors in the echinoderms

Upon determining that NGFFFamide is the endogenous ligand for an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*, NG peptide precursors and candidate NPS/CCAP-type receptors across the echinoderms were identified through both the generation and analysis of transcriptome data and from publicly available transcriptome/genome data. The purpose for the identification of NG peptide precursors and candidate NPS/CCAP-type receptors in the echinoderms is two-fold. Firstly, it enables reconstruction of the evolution of the NG peptide signalling system in deuterostomian invertebrates. Secondly, it enables prediction of the endogenous ligands for NPS/CCAP-type receptors identified in these species.

To date, a number of NG peptide precursors and NPS/CCAP-type receptors in deuterostomian invertebrates have been reported (**Table 6**). These include the SFRNGVamide precursor in the cephalochordate *B. floridae* (Elphick, 2010), the NGFWNamide/NGFYNamide precursor in the hemichordate *S. kowalevskii* (Elphick, 2010), the NGIWYamide precursor in the sea cucumber *A. japonicus* (Elphick, 2012) and the NGFFYamide precursor in the starfish *A. rubens* (see **chapter 4**). Consistent with the finding that NGFFFamide is the endogenous ligand for an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*, NPS/CCAP-type receptors are also found in species that contain an NG peptide-type precursor (Mirabeau and Joly, 2013; Pitti and Manoj, 2012; Valsalan and Manoj, 2014). Interestingly, however, to date, NG peptide precursors and NPS/CCAP-type receptors have not been identified in the urochordates, including the tunicates *Ciona intestinalis* (http://genome.jgi-psf.org/Cioin2/Cioin2.home.html) and *Oikopleura dioica* (http://www.genoscope.cns.fr/externe/GenomeBrowser/Oikopleura/). This finding suggests that the loss of a gene encoding an NG peptide precursor is

associated with a loss of an NPS/CCAP-type precursor or vice versa, providing further evidence that the NG peptides are the endogenous ligands for NPS/CCAP-type receptors.

The recent availability of echinoderm genome/transcriptome data has enabled identification of NG peptide precursors and candidate NPS/CCAP-type receptors in each of the extant classes of echinoderms (**Fig. 5.4**). The echinoderms comprise five classes – the echinoids, the holothurians, the asteroids, the ophiuroids and the crinoids (**Fig. 1.4**). The echinoids and holothurians form the echinozoan clade, the asteroids and ophiuroids form the asterozoan clade whilst the crinoids are basal to the echinozoan and asterozoan clades (O'Hara et al., 2014; Telford et al., 2014).

Species	NG peptide(s)	Ref(s)	NPS/CCAP-type receptor	Ref(s)
B. floridae	SFRNGV-NH_2 *(x1 copy)*	*a*	✓	*d*
S. kowalevskii	NGFYN-NH_2 *(x1 copy)* NGFWN-NH_2 *(x5 copies)*	*a*	✓	*d*
A. japonicus	NGIWY-NH_2 *(x5 copies)*	*b*	✓	*e*
A. rubens	NGFFY-NH_2 *(x2 copies)*	*c*	-	-

Table 6. NG peptides and candidate NPS/CCAP-type receptors previously reported in invertebrate deuterostomes. NG peptide precursors and NPS/CCAP-type receptors have been identified in the cephalochordate *B. floridae*, the hemichordate *S. kowalevskii* and the echinoderm species *A. japonicus* and *A. rubens*. *a)* Elphick, 2010; *b)* Elphick, 2012; *c)* **chapter 4**; *d)* Pitti and Manoj, 2012; *e)* Rowe et al., 2014.

5.3.3.1. NG peptide signalling system in the Echinoidea

In the echinoids, beyond the characterisation of the NG peptide signalling system in the sea urchin *S. purpuratus*, an NG peptide precursor and NPS/CCAP-type receptor were identified in the genome of the sea urchin species *L. variegatus* (http://www.echinobase.org/Echinobase/Blast/LvBlast/blast.php). In comparison with the *S. purpuratus* precursor, the *L. variegatus* precursor also comprises two copies of the NG peptide NGFFFamide in tandem flanked by dibasic cleavage sites (**Fig. 5.5A; Fig. S69**). The *L. variegatus* precursor and the *S. purpuratus* precursor both have a similar gene structure comprising three exons. In the *L. variegatus* precursor, the first exon encodes the N-terminal signal peptide and 10 amino acid residues separating the signal peptide from the NGFFFamide coding region. The second exon encodes two copies of NGFFFamide flanked by dibasic cleavage sites and a 26 amino acid sequence separating NGFFFamide from the C-terminal neurophysin domain. The third exon encodes the C-terminal neurophysin domain followed by a stop codon. Similarly, the *L. variegatus* NPS/CCAP-type receptor and the *S. purpuratus* NGFFFamide receptor also have a similar gene structure comprising nine exons (**Fig. 5.6A; Fig. S74**). This conservation of gene exon/intron structure provides further evidence of orthology.

5.3.3.2. NG peptide signalling system in the Holothurians

The holothurians (e.g. sea cucumbers) form a sister class to the echinoids (e.g. sea urchins) and together comprise the echinozoan clade (O'Hara et al., 2014; Telford et al., 2014). Previous investigation into the transcriptome of the sea cucumber *A. japonicus* (Du et al., 2012) has led to the identification of an NG peptide precursor comprising five copies of the NG peptide NGIWYamide (Elphick, 2012) and an NPS/CCAP-type receptor (Rowe et al., 2014). Interestingly, however, the *A.*

japonicus precursor does not have a C-terminal neurophysin domain, which presumably reflects loss in the holothurian lineage (Elphick, 2012). Extending analysis to the transcriptome of the sea cucumber species *P. parvimensis* reveals a precursor protein comprising four copies of the NG peptide NGIWYamide and one copy of the NG peptide NGVWYamide, differing by a single amino acid residue – a valine (V) as opposed to an isoleucine (I) (**Fig. 5.5B; Fig. S70**). The *P. parvimensis* precursor does not have a C-terminal neurophysin domain, which provides further evidence of loss of this domain in the holothurian lineage. Furthermore, an NPS/CCAP-type receptor was also identified in this species (**Fig. 5.6B; Fig. S75**).

5.3.3.3. NG peptide signalling system in the Asteroidea

The asteroids (e.g. starfish) form a sister clade to the ophiuroids (e.g. brittle stars) and together comprise the asterozoan clade (O'Hara et al., 2014; Telford et al., 2014). Analysis of transcriptome of the starfish *A. rubens* resulted in identification of an NG peptide precursor comprising two copies of the NG peptide NGFFYamide (**Fig. 2.4; Fig. S5**) and an NPS/CCAP-type receptor (**Fig. 5.6C; Fig. S6**). Analysis was extended to the genome of the starfish species *P. miniata*, which also has an NG peptide precursor comprising two copies of the NG peptide NGFFYamide (**Fig. 5.5C; Fig. S71**) and an NPS/CCAP-type receptor (**Fig. 5.6D; Fig. S76**).

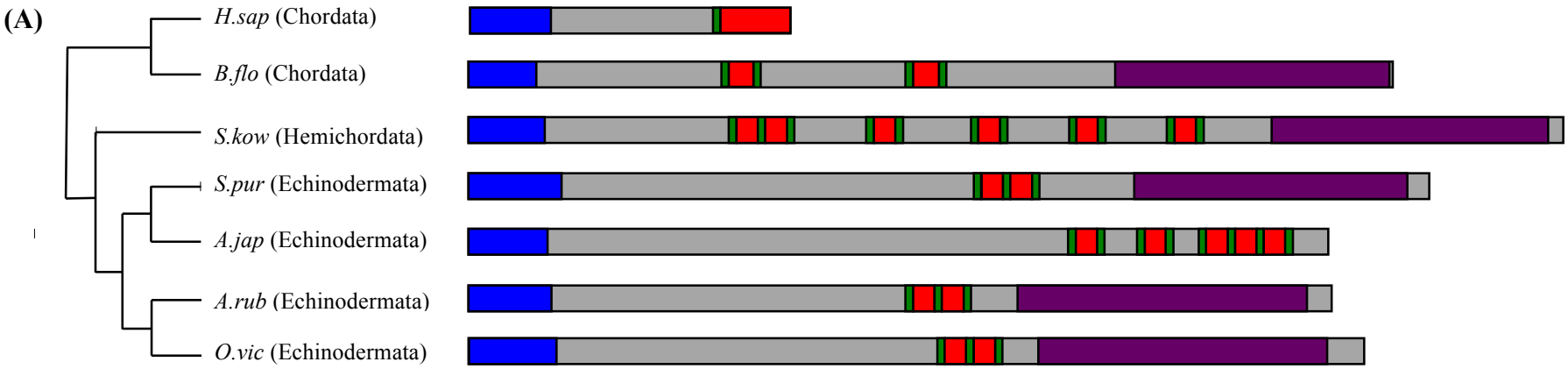
5.3.3.4. NG peptide signalling system in the Ophiuroidea

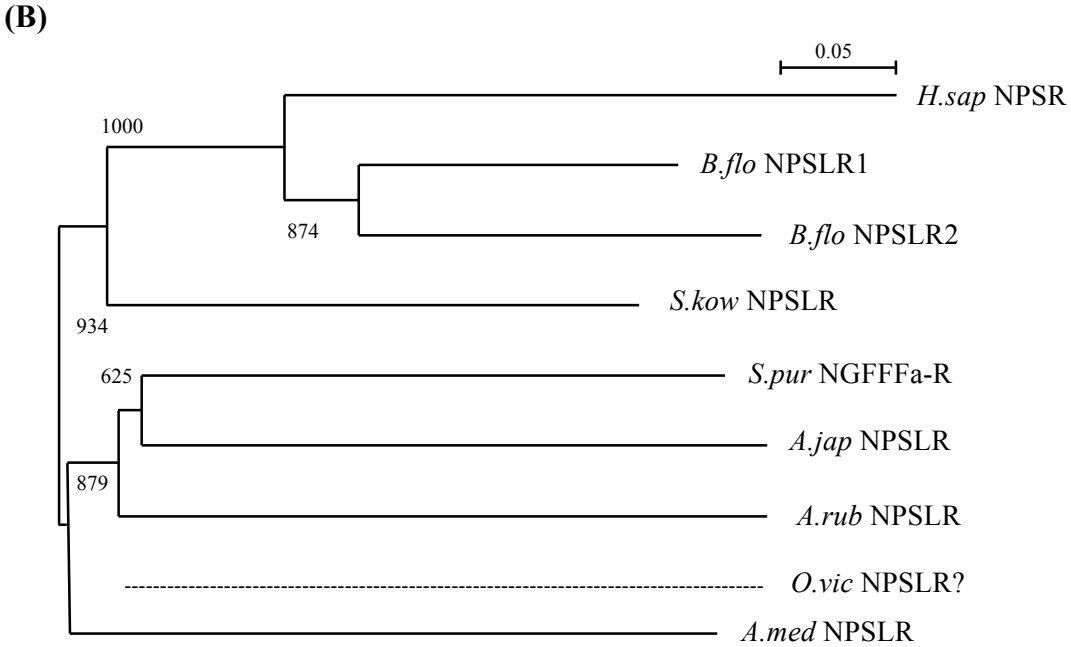
The ophiuroids (e.g. brittle stars) form the sister clade to the asteroids (e.g. starfish) comprising the asterozoan clade (O'Hara et al., 2014; Telford et al., 2014). To date, NG peptide precursors and NPS/CCAP-type receptors have not been reported for any species of ophiuroid. In this regard, a transcriptome dataset has been generated from the brittle star *O. victoriae* (Elphick et al., 2015) and a transcriptome dataset has been

analysed from the brittle star species *A. filiformis* leading to the identification of NG peptide precursors in both species. The *O. victoriae* precursor contains one copy of the NG peptide NGFFFamide and one copy of the NG peptide NGFFYamide in tandem (**Fig. 5.5D; Fig. S72**). The *A. filiformis* precursor contains two copies of the NG peptide NGFFFamide in tandem (**Fig. 5.5E; Fig. S73**). Interestingly, however, an NPS/CCAP-type receptor was not identified in either species.

5.3.3.5. NG peptide signalling system in the Crinoidea

The crinoids (e.g. feather stars and sea lilies) are basal to the echinozoan and asterozoan clades (O'Hara et al., 2014; Telford et al., 2014). To date, NG peptide precursors and NPS/CCAP-type receptors have not been reported for any species of crinoid. In this regard, a transcriptome dataset has been generated from the feather star *A. mediterranea* (Elphick et al., 2015) and a transcriptome dataset has been analysed from the feather star species *F. serratissima* (access to transcriptome dataset was provided by Dr. Christopher Lowe, Stanford University). Interestingly, through BLAST analysis, NG peptide precursors were not identified in either species. However, NPS/CCAP-type receptors were identified for both species (**Fig. 5.6E-F; Fig. S77-78**). The absence of NG peptide precursors in the crinoids could represent sequence divergence in the peptide region and/or represent the loss of the C-terminal neurophysin domain.





(C)

SFRNGVGTGMKKTSFQRAKS(x1)

SFRNGV-NH_2(x2)

NGFWN-NH_2(x5) & NGFYN-NH_2 (X1)

NGFFF-NH_2(x2)

NGIWY-NH_2(x5)

NGFFY-NH_2(x2)

NGFFF-NH_2(x1) & NGFFY-NH_2 (x1)

?

Fig. 5.4. The NPS/NG peptide precursors and receptors in the deuterostomes. **(A)** Schematic depicting the structural organisation of NPS/NG peptide precursors in the deuterostomes. Signal peptides are represented in blue, putative neuropeptides are represented in red, dibasic cleavage sites are represented in green and C-terminal neurophysin domains are represented in purple. **(B)**. Neighbour joining (NJ) tree (including bootstrap values out of 1000) of NPS/CCAP-type receptors in the deuterostomes. The hypothetical position of a candidate NPS/CCAP-type receptor in the brittle star *O. victoriae* is depicted by a dashed line. **(C)** Alignments of putative neuropeptides derived from NPS/NG peptide precursors. Conserved NG motifs are highlighted in yellow and the number of copies of each peptide is represented in parentheses. *A.jap*, *Apostichopus japonicus*; *A.med*, *Antedon mediterranea*; *A.rub*, *Asterias rubens*; *B.flo*, *Branchiostoma floridae; H.sap*, *Homo sapiens*; *O.vic*, *Ophionotus victoriae*; *S.kow*, *Saccoglossus kowalevskii*; *S.pur*, *Strongylocentrotus purpuratus*.

(A) *L. variegatus* NGFFFamide precursor

MGYERRILRTLLSISIVLASFVTVYGERDNNFMQQKQFRNIVPSPLIQKWRENQMGP
SEEKISNEEWKDELLSNLRNVLRKHNPSSRSRDRTDITAYGLQEPLPQLPADVSADQ
LFILEGASPNSPLGNYEAETPIDEDKRNGFFFGKRNGFFFGKRSGSEVSATKVDDDR
LSKFESSGSFDKCRPCGPRSQGRCVMVGTCCSPQFGCYLFTPEAAACMTEDVSPCQL
NAPSCGLEGQCVANGICCSAAEGACHLDPTCTSMSLK

(B) *P. parvimensis* NGIWY/NGVWYamide precursor

MAVEAKIVSCLVCIWLTSTVYSQSNTGRTHDYGELSKAVDTFLDILMDEENFDDVNN
IESWETVLKEDINPKLRILAHVMRSLSSRQDTSSLRQQVLPNDYISRYLQEFLTDEQ
PFWDESSPKLPSLQTPELDQIKASADERNNFWSDNPSHRPPEGIPFSAGEESKRNGI
WYGKRKSSLQDGEAVKRNGVWYGKRSSPPVDDKRNGIWYGKRNGIWYGKRNGIWYGK
RDDSLYSEEMM

(C) *P. miniata* NGFFYamide precursor

MGGRSVSLIGVVIVVVTILVAFTWAETANGSLSQKVKIRREGRESDKYWSKDVGLTD
QQLRQLLANSLTNSYSTGEGESQTRPAKPGYRLYGNRLSDTNLGIDGNEEDEVETEV
TPRGSSLRTSKRNGFFYGKRNGFFYGKRSGSTLENANQCTQCGPQNSGQCVMFGTCC
SYAAGGCFFMTEEALPCVTSKTSSICQLQGAPCGSDGYGRCVADSVCCSPQEGACHV
DVACSSKPTYQ

(D) *O. victoriae* NGFFF/NGFFYamide precursor

MAVGIRNVILHLILVLYVARTILGEIDTYKVRRYTDGSPVNSNHWTQSDNINELRKE
IFASLPADLPGIIHKSKTHDSRSGRHSDELSNLEQLGNVAGKDGANDAKDKLALYNF
MSKQATRNNLGDELDKRNGFFFGKRNGFFYGKRDVEAVNNEDSCVRCGPENRGQCVM
FGTCCSPQFGCYLMTKESEACMSNHIGTGCRNLDMAPQCGSTGVCVAKGVCCSPQDG
ACHIDVTSCLSDNTTIDKSL

(E) Partial (5') *A. filiformis* NGFFFamide precursor

DIPAILLNPQTQENKEDTPQDLTALEQVERVATGQTIDTDARDNLYNFLSRQSARNN
YQAGLESKRNGFFFGKRNGFFFGKRDTEAVKSDLCTSCGPQNIGQCVMFGTCCSPQF
GCYFMTQEATACTKHHIDNACWNQDLMPSCGRRGICAADALCCSPKDGACRIDLSCS
SPQKKNYDFND

Fig. 5.5. NG peptide precursors identified in echinoderm species. (A) *L. variegatus* NGFFFamide precursor. **(B)** *P. parvimensis* NGIWY/NGVWYamide precursor. **(C)** *P. miniata* NGFFYamide precursor. **(D)** *O. victoriae* NGFFF/NGFFYamide precursor. **(E)** Partial (5') *A. filiformis* NGFFFamide precursor. N-terminal signal peptides are represented in blue, putative NG peptides are represented in red, C-terminal glycine (G) residues are represented in orange and dibasic cleavage sites are represented in green. Neurophysin domains characteristic of NG peptide precursors are represented purple, with the characteristic fourteen cysteine (C) residues underlined.

(A) *L. variegatus* NPS/CCAP-type receptor

MTTQLNFDPEVMTTEALNFTDSWPENDTGNGIVDRWSLDKHIQLAILWILFTLIIVG
NGIVLIALWLVRHKKSRLNFFITNLAVADICVGLFSVGLDIIDRQTPEFTGGDIACK
LYRYVQAYVVLASSYQLVALSFDRFFAIVFPMDFTGNHAGKRSTLLAAGGWLLPSVL
CITSAVVFQVDLLDSPDRKSKVMSCWPAPLYGDQAWKLKVYSMYVISSFFYIPLILI
TFCYVTIIVTIWKRAKKMGGPQVIKKSKNSNRDVAYEGLSKDSNCTMPKHRASSRGL
IPKAKVKTIKMTICIVLAYICCFLPFSLFYTLEAFGCINPQSTAVLMATPVLQNLPS
LNSATNPFIYGIFSTNVCKELRRIPAINWIADKVPCCTAWKPLRFGRTTYQTNTHTT
EFNNFSDGHTGSRGRNIVSMSGKVVDGPSRDDSRNSTTSPM

(B) *P. parvimensis* NPS/CCAP-type receptor

MADEPAVLSTLGTELQRTSENTFTTIGSHEHDNRLAVESELQLAILWTMYILIVVGN
GLVLLALFSVRHKKSRLNFFVKHLAIADVCVGLLNVLPEIIHRYRGAFYAGMFLCKI
KSYGQAFVIYASIYQMVALSLDRFFAIVFPMDFMASRKRSTFMAAGAWILPGMLATP
SLAIFVTVDLHGQTQCAPIAILQDKLKYQLYSLYIVSITFLIPLMILCVCYGTMISV
IWRRGKAMAPPVKSDKNANSGGVKYTGLQKKAKQDENNFKHRSSSRGLIPRAKIKTV
KMTICIVIAYILCWLPTSLYFTLEAFKVVKPSADPQHAIYWVSVIMQNLVYLNSATN
PFIYGFFSSNICKELRRYYIIRQLLKWMPCCKVTEPGYGRSTAGTVMTEFHSHTAAI
SDNHRFNHNASSNDEKSVRETSHI

(C) *A. rubens* NPS/CCAP-type receptor

MATIPAYDHLVTDSVMAGYSLNDTASTVMVPTGLPSTLEGAPNATTSVTYFSDGENR
LSFYGGFQLIVLWVLFGLTVIGNSTVLLAVYVIRHKKSRLNFFVAHLAASDLLVGIV
NNGYEVLYRYLGEFYGGMVFCKIIRFSQAYVINASSFQLVALSLDRFFAIVFPMDFS
GSGKRANLMAVTAWIAPLFASIPSAIVFEAAVDSWGKTHCIPPPLVPGSWQYKVYTL
YVVSGFFYIPLIIISTCYIFMVISIWRRSKYMMGQKPEKVKGKAAKNASKEKEPMKH
RASSRGLIPKAKIKTLKMTVSIIVAFIVCWCPFSVFYTLDAFGVIFIDEANLNTAFR
ASAFIQNLPFLNSAINPFIYGMFSTNICQELRRFSVINWMATKLRCCKSWRPSRYGR
STTLRTDTNLTDMSEGASGTHRGHTRPLYVYTPGRTSNSDHSRDLSNSAM

(D) *P. miniata* NPS/CCAP-type receptor

MTSLATSTSYERLMNYSSGGGVSVMANPAGVDLSTEGALNASTTENPFADGENRISF
YGDLQLIVLWVLVALTVVGNSTVLLAVFVIRHKKSRLNFFVAHLAAADLLVGVINAG
YEIIYKYMGEFYGGTVFCKIIRYLQAYVINASSFQLVALSLDRFFAIVFPMDFSGSG
KRANLMAGTAWIAPFFASIPSAIVFEAKMDGWGTTHCYPPPLQPGSWQYKVYTLYVV
SGFFYIPLIVISTCYIFMVVSIWRRSKYMMGQKPDKVKGKAASKEMKHRASSRGLIP
KAKIKTLKMTVCIIVAFILCWCPFSIFYTLDAFRVIVIDEATLNTAFRASAFIQNLP
YLSSAINPFIYGLFSTNICQELRRFRVINWMATKLRCCKSWRPSRYGRSTLRTDTNL
TDMSEGASGTHRGHSRPMYVYTPARTSNSDNSREMPNSAM

(E) *A. mediterranea* NPS/CCAP-type receptor

MDYNMDNYTDFWTEMTPTTKEVSRIVLYRDLQLIVLWSLFISILIGNGAVLFALYTV
RHKKSRMNFFVMNLAISDIIVGFFEVLVQLIHRGYTGTWETGNIACKLVKYIQAASL
TASSCQLVALSVDRYLAIIYPMNFSGRSRRAHMMAAAAWITPFITSITSLHIFEIRE
HGEGTQCWMVLENENLKLKLYTIYVMCILFFIPLVVILFCYCSIIITIWKKSKMMGP
NIKPASNKKGNNYDNLKADEGSRSHRASSRGLIPKAKVKTIYITLSIVIAFVLCWSP
FIFTYLLAAFQVIQPSAKLMAVIANLPAINSAVNPLIYGIFSTNLCRELKRIPVINW
FAGVLPCCTARKKAEPGFTRAMYTRTENTNMDYSVTTCEQDSERIKLGQTSK

(F) ***F. serratissima*** **NPS/CCAP-type receptor**

MDDMDNTTEFWSEMTPTLEVSRIVLHRDLQLVVLWSLFVSIFIGNGAVLLALYSVRH
KKSRMNFFVMNLAISDILVGIFEVLVQLIHRGYTGTWRTGNISCKIVKYIQAASMSA
SSFQLVALSVDRYLAIIYPMNFSGRSRRAHIMAAAAWIAPFITSITSLHIFEIREFP
EGPQCWMTLEGGGWKSKLYTVFVCCILFFIPLIVILFCYCSIIITIWQKSKMMGPKI
KSTNNKKGNYDNLKADEGSRSHRASSRGLIPKAKIKTIYITLSIVIAFVLCWSPFIF
YYLLMAFKVIVPDPKVMAVIANLPAINSAVNPLIYGIFSTNLCRELK

Fig. 5.6. NPS/CCAP-type receptors identified in echinoderm species. (A) *L. variegatus* NPS/CCAP-type receptor. (**B)** *P. parvimensis* NPS/CCAP-type receptor. (**C)** *A. rubens* NPS/CCAP-type receptor. (**D)** *P. miniata* NPS/CCAP-type receptor. (**E)** *A. mediterranea* NPS/CCAP-type receptor. (**F)** *F. serratissima* NPS/CCAP-type receptor. Predicted transmembrane domains are highlighted in grey.

5.4. Discussion

The finding that the NG peptide NGFFFamide activates an NPS/CCAP-type receptor in the sea urchin *S. purpuratus*, which to date is the first echinoderm neuropeptide GPCR to be deorphanised, has unified a bilaterian family of neuropeptides that encompasses NPS-type peptides in tetrapod vertebrates, NG peptides in deuterostomian invertebrates and CCAP-type peptides in protostomian invertebrates. This finding, along with the identification of novel NG peptide precursors and/or NPS/CCAP-type receptors across each of the five extant echinoderm classes (echinoids, holothurians, asteroids, ophiuroids and crinoids), has allowed reconstruction of the evolution of the NPS/NG peptide/CCAP-type neuropeptide signalling system.

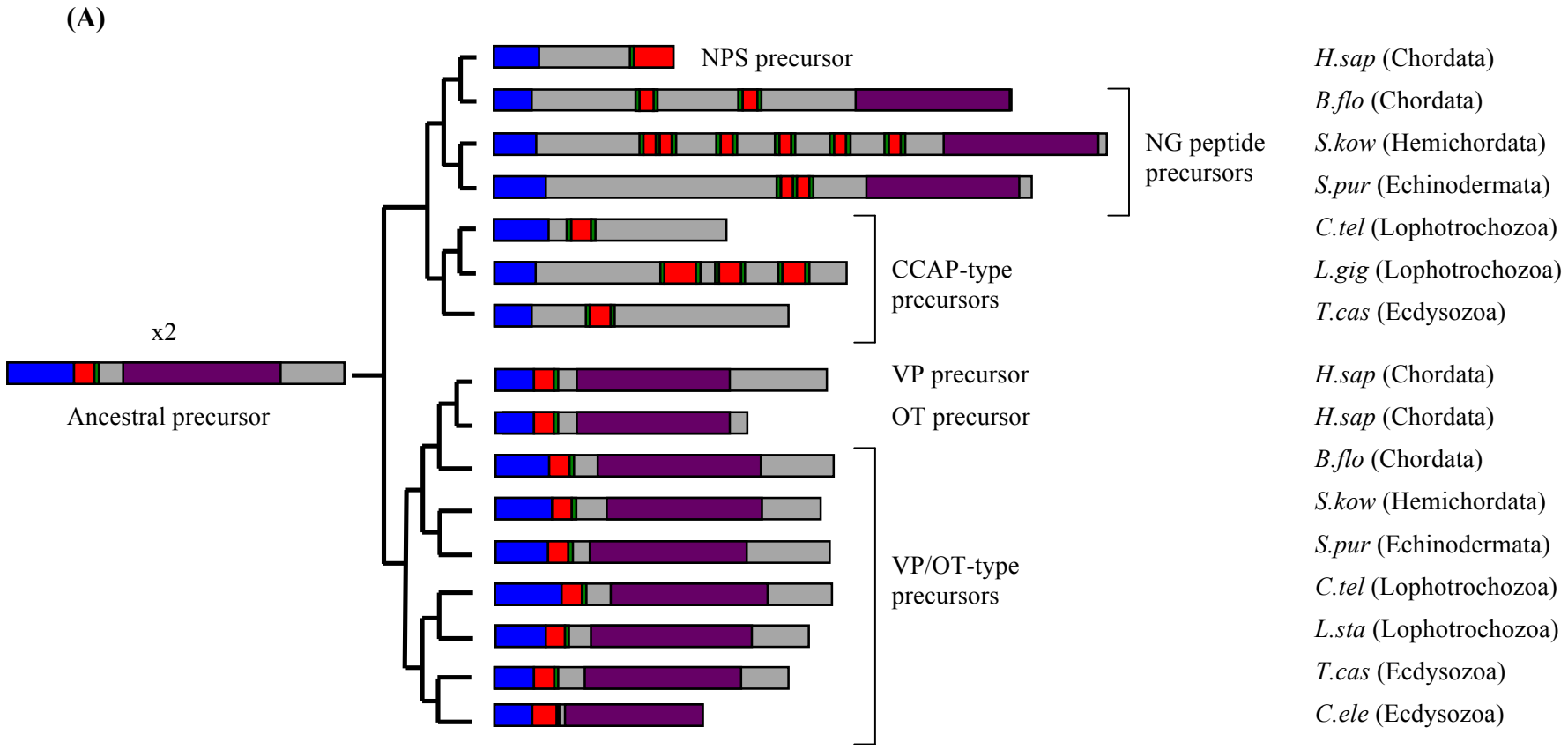
5.4.1. An ancient duplication of a VP/OT-type neuropeptide signalling system gave rise to the NPS/NG/CCAP-type signalling systems

In **Fig. 5.7**, an evolutionary hypothesis is presented in which an ancient duplication of a VP/OT-type precursor and receptor in the common ancestor of the Bilateria gave rise to the VP/OT-type neuropeptide signalling system, the NPS-type system in tetrapod vertebrates, the NG peptide system in deuterostomian invertebrates and the CCAP-type system in protostomian invertebrates. In the common ancestor of the Bilateria, one copy was retained giving rise to the highly conserved VP/OT-type precursors found throughout the extant bilaterians, comprising a single VP/OT-type peptide with a disulphide bridge and a C-terminal neurophysin domain. The second copy diverged but this took different courses during the evolution of protostomes, deuterostomian invertebrates and vertebrates. In this regard, comparison of the sequences of NPS-type peptides, NG peptides and CCAP-type peptides with each

other and with the sequences of VP/OT-type peptides show the differential retention of a number of structural characteristics, which reflect the common evolutionary origins of these neuropeptides.

In the deuterostomian branch of the animal kingdom, the second copy lost the disulphide bridge characteristic of VP/OT-type peptides but differentially retained the C-terminal neurophysin domain. In the deuterostomian invertebrates, the NG peptide family retained the C-terminal neurophysin domain, with the exception of the holothurian lineage (e.g. sea cucumbers). However, in the vertebrate lineage that gave rise to NPS-type precursors in tetrapod vertebrates, the C-terminal neurophysin domain was lost. To date, NPS-type precursors have not been identified in the teleosts and basal vertebrate classes including the Chondrichthyes (e.g. cartilaginous fish) and Agnathans (e.g. jawless fish) (Reinscheid, 2007). This indicates that there has been independent loss of NPS-type precursor genes in multiple vertebrate lineages whilst there has also been loss in the urochordates.

In the protostomian branch of the animal kingdom, the second copy retained the disulphide bridge characteristic of VP/OT-type peptides. Interestingly, however, comparison with the NG peptides and NPS-type peptides reveals that the NG motif is also a feature of CCAP-type peptides in molluscs (e.g. *Conus villepinii*; [GI: 325529921] and *Aplysia californica*; [GI: 524893759]). It is therefore possible that the NG motif may be a unifying characteristic of this bilaterian peptide family. However, CCAP-type precursors did not retain the C-terminal neurophysin domain characteristic of VP/OT-type precursors suggesting that this domain was lost during the evolutionary history of CCAP-type precursor genes.


(A)
x2
Ancestral precursor
NPS precursor
NG peptide precursors
CCAP-type precursors
VP precursor
OT precursor
VP/OT-type precursors
H.sap (Chordata)
B.flo (Chordata)
S.kow (Hemichordata)
S.pur (Echinodermata)
C.tel (Lophotrochozoa)
L.gig (Lophotrochozoa)
T.cas (Ecdysozoa)
H.sap (Chordata)
H.sap (Chordata)
B.flo (Chordata)
S.kow (Hemichordata)
S.pur (Echinodermata)
C.tel (Lophotrochozoa)
L.sta (Lophotrochozoa)
T.cas (Ecdysozoa)
C.ele (Ecdysozoa)

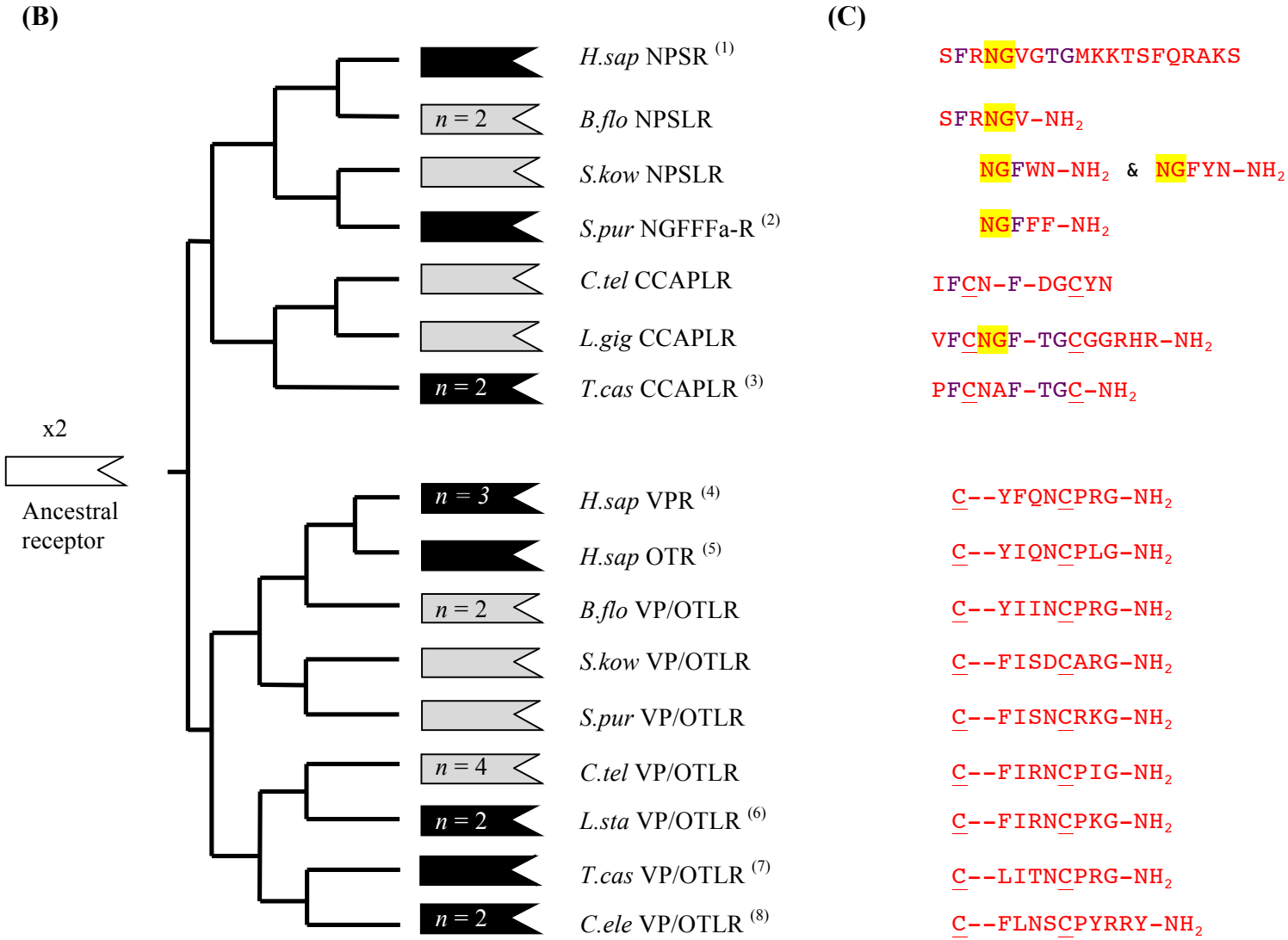

(B)
(C)
x2
Ancestral receptor
n = 2
n = 3
n = 4
H.sap NPSR (1)
B.flo NPSLR
S.kow NPSLR
S.pur NGFFFa-R (2)
C.tel CCAPLR
L.gig CCAPLR
T.cas CCAPLR (3)
H.sap VPR (4)
H.sap OTR (5)
B.flo VP/OTLR
S.kow VP/OTLR
S.pur VP/OTLR
C.tel VP/OTLR
L.sta VP/OTLR (6)
T.cas VP/OTLR (7)
C.ele VP/OTLR (8)
SFRNGVGTGMKKTSFQRAKS
SFRNGV-NH2
NGFWN-NH2 & NGFYN-NH2
NGFFF-NH2
IFCN-F-DGCYN
VFCNGF-TGCGGRHR-NH2
PFCNAF-TGC-NH2
C--YFQNCPRG-NH2
C--YIQNCPLG-NH2
C--YIINCPRG-NH2
C--FISDCARG-NH2
C--FISNCRKG-NH2
C--FIRNCPIG-NH2
C--FIRNCPKG-NH2
C--LITNCPRG-NH2
C--FLNSCPYRRY-NH2

Fig. 5.7. Evolution of the VP/OT-type and the NPS/NG/CCAP-type neuropeptide signalling systems in the Bilateria. (A) Schematic depicting an ancient duplication of a VP/OT-type precursor in the common ancestor of the Bilateria that gave rise to the highly conserved VP/OT-type precursors in extant bilaterian animals and the divergent NPS, NG peptide and CCAP-type precursors in tetrapod vertebrates, deuterostomian invertebrates and protostomian invertebrates respectively. Signal peptides are represented in blue, putative neuropeptides are represented in red, dibasic cleavage sites are represented in green and C-terminal neurophysin domains are represented in purple. **(B)** Schematic depicting an ancient duplication of a VP/OT-type receptor in the common ancestor of the Bilateria that gave rise to NPS/NG/CCAP-type receptors. Receptors for which the cognate ligand has been proven are highlighted in black whilst receptors for which the cognate ligand has yet to be proven are highlighted in grey. Multiple isoforms of receptors are represented by *n* = x. **(C)** Alignment of NPS/NG/CCAP-type peptides and VP/OT-type peptides derived from precursors shown in **(A)** that are proven or candidate ligands for the receptors shown in **(B)**. NG motifs are highlighted in yellow and pairs of cysteine (C) residues that form disulphide bridges in VP/OT-type and CCAP-type precursors are underlined. A TG motif and two phenylalanine (F) residues, which are conserved in some peptides, are highlighted in purple. *B.flo*, *Branchiostoma floridae*; *C.ele, Caenorhabditis elegans; C.tel, Capitella teleta; H.sap*, *Homo sapiens*; *L.gig, Lottia gigantea; L.sta, Lymnaea stagnalis; S.kow*, *Saccoglossus kowalevskii*; *S.pur*, *Strongylocentrotus purpuratus*; *T.cas, Tribolium castaneum. 1*, Xu et al., 2004; *2*, this **chapter**; *3*, Stafflinger et al., 2008; *4*, Birnbaumer et al., 1992; Sugimoto et al., 1994; Thibonnier et al., 1994; *5*, Kimura et al., 1992; *6*, van Kesteren et al., 1995; *7*, Li et al., 2011; *8*, Beets et al., 2012; Garrison et al., 2012.

Importantly, in addition to identifying NGFFFamide as the endogenous ligand for the NPS/CCAP-type receptor in the sea urchin *S. purpuratus*, there are additional pieces of evidence to support this evolutionary hypothesis including gene contiguity in the genomes of the deuterostomian invertebrates *S. purpuratus* and *B. floridae*. In the echinoderm *S. purpuratus*, a gene encoding a VP/OT-type receptor (SPU_021290) is located adjacent to a gene encoding a candidate NPS/CCAP-type receptor (SPU_021291) (**Fig. 5.1**); in the cephalochordate *B. floridae,* a gene encoding a VP/OT-type neuropeptide precursor (Brafl1-84802) is located adjacent to a gene encoding the SFRNGVamide neuropeptide precursor (Brafl1-84803) (Mirabeau and Joly, 2013). The gene contiguity in the genomes of *S. purpuratus* and *B. floridae* presumably reflects conservation of the adjacent positioning of genes generated by a tandem gene duplication event that occurred in the common ancestor of the Bilateria.

5.4.2. Characteristics of an ancestral neuropeptide precursor protein

It is important to recognise that alternative evolutionary scenarios of this bilaterian neuropeptide family have also been postulated with regards to the nature of the ancestral neuropeptide precursor that gave rise to NPS-type peptides in tetrapod vertebrates, NG peptides in the deuterostomian invertebrates and CCAP-type peptides in the protostomian invertebrates (Mirabeau and Joly, 2013; Valsalan and Manoj, 2014). It has been proposed that the ancestral precursor comprised both an NPS-type and VP/OT-type peptide associated with a C-terminal neurophysin domain before duplication in the common ancestor of the bilaterians (Mirabeau and Joly, 2013; Valsalan and Manoj, 2014) (**Fig. 5.8**).

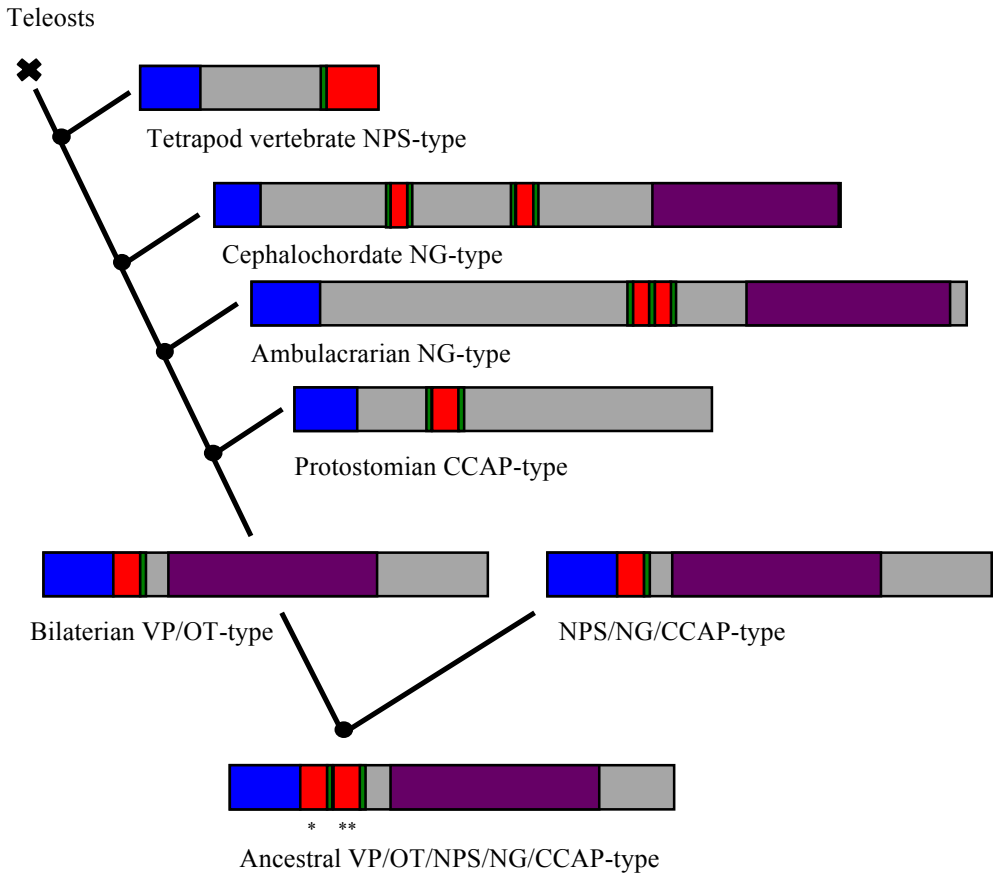


Fig. 5.8. Characteristics of a common ancestral neuropeptide precursor protein. Mirabeau and Joly, 2013 propose that the ancestral precursor protein comprised both a VP/OT-type and an NPS/NG/CCAP-type peptide before subsequent duplication giving rise to VP/OT-type precursors in extant bilaterians and the divergence of the second copy into the NPS/NG/CCAP-type neuropeptide signalling systems. Signal peptides are represented in blue, putative neuropeptides are represented in red, dibasic cleavage sites are represented in green and C-terminal neurophysin domains are represented in purple. An asterisk denotes a VP/OT-type peptide whilst a double asterisk denotes an NPS/NG/CCAP-type peptide in the ancestral precursor. Figure adapted from Mirabeau and Joly, 2013.

However, a scenario in which the ancestral precursor encoded a single VP/OT-type peptide before duplication in the common ancestor of the Bilateria and subsequent divergence of the second copy to give rise to NPS in tetrapod vertebrates, NG peptides in deuterostomian invertebrates and CCAP-type peptides in protostomian invertebrates seems more likely. This evolutionary scenario differs from previously reported evolutionary hypotheses on the basis of two factors. The first is that VP/OT-type peptides are highly conserved in extant animals throughout the animal kingdom as opposed to NPS-type peptides, NG peptides and CCAP-type peptides; this alludes to the VP/OT-type peptides reflecting the ancestral state. The second factor is the interactive stoichiometry between VP/OT-type peptides and the neurophysin domain, which has been shown to be 1:1 (de Bree, 2000; de Bree and Burbach, 1998). However, the NPS/NG peptide/CCAP-type precursors are highly divergent in terms of the number of neuropeptide copies and the presence of a C-terminal neurophysin domain. For example, the neurophysin domain has been lost in NPS-type and CCAP-type precursors whilst NG peptide precursors in deuterostomian invertebrates comprise two to six NG peptides. Therefore, with the 1:1 interactive stoichiometry of VP/OT-type peptides with the neurophysin domain, a hypothesis in which the ancestral precursor comprised a single VP/OT-type peptide is favoured over the alternative model where the ancestral precursor comprised a VP/OT-type and a NPS/NG peptide/CCAP-type peptide (**Fig. 5.8**) (Mirabeau and Joly, 2013; Valsalan and Manoj, 2014).

An interesting question that remains unanswered is the significance of neurophysin domain retention in some, but not all, NG peptide precursors in the deuterostomian invertebrates, with the neurophysin domain appearing to be lost in the holothurian lineage. It remains unclear whether the neurophysin domain is required for biosynthesis of the NG peptides as is observed for VP/OT-type peptides

(Cool et al., 2008; de Bree, 2000; de Bree and Burbach, 1998). If it is required, what is the interactive stoichiometry? If it is not required, what is the function of the neurophysin domain? Investigation into the functional roles of the neurophysin domain in NG peptide precursors will shed further light on the evolutionary origins of the NG peptide family and the composition of an ancestral precursor in the common ancestor of the Bilateria.

5.4.3. An evolutionarily ancient role for the NPS/NG/CCAP-type neuropeptide signalling system?

The finding that the NG peptide NGFFFamide activates an NPS/CCAP-type receptor in the sea urchin *S. purpuratus* provides a foundation for investigating the physiological roles of NG peptides in deuterostomian invertebrates and has the potential to shed light on the evolution of the physiological functions of NPS/NG peptide/CCAP-type neuropeptides in the Bilateria. To date, investigations into the NG peptides in the deuterostomian invertebrates have focused on the echinoderms. For example, the NG peptide NGIWYamide in the sea cucumber *A. japonicus* has been shown to be involved in reproduction inducing *in vitro* germinal vesicle breakdown and oocyte ovulation (Kato et al., 2009), whilst the NG peptide NGFFYamide in the starfish *A. rubens* has been shown to be involved in cardiac stomach contraction and retraction (see **chapter 4**). However, to date, the roles of NG peptides in the cephalochordates (e.g. *B. floridae*) and the hemichordates (e.g. *S. kowalevskii*) remain unknown. NPS stimulates arousal and inhibits anxiolytic-like behaviour in rodents (Xu et al., 2004), with polymorphisms in the NPSR also associated with panic disorders in humans (Domschke et al., 2011; Okamura et al., 2007). CCAP, on the other hand, has been shown to stimulate cardiac activity

(Stangier et al., 1987) and trigger the motor programme that enables shedding of the exoskeleton during ecdysis (Park et al., 2003) in arthropods.

Based upon the known physiological functions of NPS-type, NG peptides and CCAP-type peptides to date, it appears that this newly unified family of peptides may have an important role in triggering behaviours associated with heightened states of arousal. It will be interesting to continue investigations into the physiological roles of these peptides in a range of phyla across the animal kingdom to begin to reconstruct how this highly divergent peptide family acquired distinct physiological roles following ancient gene duplication from a VP/OT-type peptide in the common ancestor of the Bilateria.

5.4.4. Future directions

The use of genome/transcriptome data from an echinoderm species in the unification of a bilaterian neuropeptide family highlights the general importance of the deuterostomian invertebrates, and the echinoderms in particular, in bridging the gap of evolutionary understanding between neuropeptide signalling systems in vertebrates and classic model protostomian invertebrates (e.g. *Drosophila* and *C. elegans*). Moreover, the characterisation of the NGFFFamide receptor could potentially lead to synergistic exchange between research in the fields of NPS-type peptides, the NG peptides and CCAP-type peptides. For example, research in the field of the NG peptides and CCAP-type peptides may inform the pharmaceutical development of synthetic NPS receptor agonists that could be used as anxiolytics. Conversely, NPS receptor antagonists may inform the pharmaceutical development of CCAP antagonists that could be used for pest control through the inhibition of ecdysis.

6. General discussion

Neuropeptides are evolutionarily ancient mediators of neuronal signalling that regulate a diverse range of physiological processes and behaviours across the animal kingdom. Recent advances in comparative genomics/transcriptomics has provided new opportunities to analyse neuropeptide systems in a wider range of phyla – from model protostomian invertebrates including *Drosophila melanogaster* (Hewes and Taghert, 2001) and *Caenorhabditis elegans* (Nathoo et al., 2001) to deuterostomes including the urochordate *Ciona intestinalis* (Kamesh et al., 2008) and the cephalochordate *Branchiostoma floridae* (Nordstrom et al., 2008). It has therefore become possible, in many cases, to trace the evolutionarily origins of neuropeptide signalling systems back to the common ancestor of the Bilateria (Jekely, 2013; Mirabeau and Joly, 2013).

6.1. *A. rubens*: providing novel insights into neuropeptide evolution

The echinoderms are of particular interest from an evolutionary perspective because as deuterostomian invertebrates they occupy an intermediate position in animal phylogeny, bridging from model protostomian invertebrates (e.g. *Drosophila* and *C. elegans*) to the vertebrates. The generation and analysis of genome/transcriptome data from the sea urchin *Strongylocentrotus purpuratus* (Burke et al., 2006; Rowe and Elphick, 2012), the sea cucumber *Apostichopus japonicus* (Rowe et al., 2014) and the starfish species *Patiria miniata* (Elphick et al., 2013) have provided important insights into neuropeptide diversity and its functional significance across the animal kingdom.

The common European starfish *Asterias rubens* has been used as a model system for neuropeptide research for over twenty five years, with the SALMFamides

S1 and S2 the first neuropeptides to be identified in an echinoderm species (Elphick et al., 1991a; Elphick et al., 1991b). However, with no genome or transcriptome data available, the sequencing of a neural transcriptome of the starfish *A. rubens* has provided an opportunity to identify novel neuropeptide signalling systems in the species and has provided a foundation to investigate neuropeptide diversity and its functional significance across the animal kingdom.

The analysis of the neural transcriptome of *A. rubens* has revealed the presence of thirty-five novel neuropeptide precursors and a number of candidate receptors (see **chapter 2**). Interestingly, precursors of two kisspeptin (KP)-type peptides (the first to be identified in a non-chordate species), a melanin-concentrating hormone (MCH)-type peptide (the first to be identified outside of the vertebrates) and two tachykinin (TK)-type peptides (the first to be identified in an ambulacrarian species) have been identified. However, before undertaking functional studies on these novel neuropeptides, it will be important to confirm predicted peptide sequences and potential post-translational modifications through mass spectrometry.

6.2. *A. rubens*: providing novel insights into neuropeptide function

The utilisation of the *A. rubens* neural transcriptome and subsequent sequence confirmation through mass spectrometry has allowed for functional studies on two novel neuropeptides – a vasopressin/oxytocin (VP/OT)-type peptide ("asterotocin") (see **chapter 3**) and an NG peptide ("NGFFYamide") (see **chapter 4**). Investigation into the physiological functions of asterotocin and NGFFYamide has revealed that asterotocin causes cardiac stomach eversion *in vivo* and cardiac stomach relaxation *in vitro* whilst the NG peptide NGFFYamide causes cardiac stomach retraction *in vivo* and cardiac stomach contraction *in vitro*.

The discovery that asterotocin and NGFFYamide are involved in the

processes of cardiac stomach eversion and retraction respectively, crucial to the extraoral feeding process in the starfish *A. rubens* (see **chapter 1.3.2.3)**, has provided insights into the evolutionarily conserved roles of the VP/OT-type and NG peptide signalling systems (see **chapter 5**). However, understanding the function of neuropeptides that control feeding in starfish will also have both economic and environmental implications. For example, some starfish species, including *A. rubens,* feed on economically important shellfish (Agüera et al., 2012; Dare, 1982; Dolmer, 1998; Magnesen and Redmond, 2012), whilst other starfish species including the crown-of-thorns starfish *Acanthaster planci* feed on reef-building corals and cause destruction of Pacific reef tracts (De'ath et al., 2012; Kayal et al., 2012; Timmers et al., 2012). Thus, the identification of neuropeptides that regulate cardiac stomach eversion (asterotocin) or retraction (NGFFYamide) may provide a basis for development of non-peptidic small molecule agonists and antagonists that mimic or block the effects of these endogenous neuropeptides; this could have implications for the chemical control of starfish feeding.

6.3. The echinoderms: providing novel insights on neuropeptide evolution

The sequencing of a neural transcriptome of the starfish *A. rubens* has provided further novel insights into the evolutionary origins of a number of neuropeptide signalling systems (see **chapter 2**). However, the genome of the sea urchin *S. purpuratus* has also proved an invaluable tool in understanding the evolution of neuropeptide signalling systems. The identification of the receptor for the NG peptide NGFFFamide – a neuropeptide-S/crustacean cardioactive peptide (NPS/CCAP)-type receptor – has unified a bilaterian neuropeptide family that includes NPS-type peptides in tetrapod vertebrates, NG peptides in deuterostomian invertebrates and CCAP-type peptides in protostomian invertebrates (see **chapter 5**).

The utilisation of genome data from the sea urchin *S. purpuratus* and genome/transcriptome data from a number of other echinoderm species has also provided a foundation to characterise candidate NG peptide receptors across the echinoderms (see **chapter 5**). This could be particularly important for the NG peptide signalling system in the sea cucumber *A. japonicus*; the NG peptide NGIWYamide induces oocyte maturation and gamete spawning (Kato et al., 2009) and could therefore have implications for sea cucumber aquaculture (Anderson et al., 2011).

6.4. Future directions: a platform for neuropeptide research

The sequencing of a neural transcriptome of the starfish *A. rubens* has provided important insights into the neuropeptide diversity and its functional significance across the animal kingdom – the use of this transcriptome data can be highlighted by functional studies on asterotocin (see **chapter 3**) and NGFFYamide (see **chapter 4**). The identification of a total of thirty-five neuropeptide precursors have also provided a foundation to investigate the function of other neuropeptide signalling systems in the starfish *A. rubens*.

However, perhaps one of the most interesting discoveries is the identification of a neuropeptide precursor that comprises two KP-type peptides – the first KP-type peptides to be identified in a non-chordate species (see **chapter 2.3.2.10**). In mammals, KP has a key role in regulating reproductive functions in the hypothalamic-pituitary axis (Oakley et al., 2009) through the regulation of gonadotropin-releasing hormone (GnRH) neurons both directly (Colledge, 2009) and indirectly (Pielecka-Fortuna et al., 2008; Zhang et al., 2008). Accordingly, mutations in the KP receptor (GPR54) have been shown to cause delayed puberty in humans (Seminara et al., 2003), GPR54-knockout mice (Funes et al., 2003) and KP-knockout

mice (d'Anglemont de Tassigny et al., 2007). Investigating KP-type neuropeptide signalling systems outside of the vertebrates could therefore be of interest both from an evolutionarily perspective and from a clinical perspective. For example, salmon calcitonin (CT) (Niall et al., 1969) is more potent in humans than human CT and has been used for the treatment of hypercalcemia and osteoporosis (Torres-Lugo and Peppas, 2000).

Informed by mRNA *in situ* hybridisation (ISH) and immunocytochemical (ICC) studies, the potential physiological roles of KP-type peptides as an evolutionarily ancient regulator of GnRH release and reproductive processes could then be investigated. In the starfish *A. rubens*, gonadal growth occurs through the autumn and winter, with gonadal size and oocyte diameter peaking in early spring prior to spawning (Barker and Nichols, 1983). It will be of interest to determine if there are seasonal changes in KP expression (and subsequent GnRH expression) through the employment of radioimmunoassays (RIA) and the quantitative polymerase chain reaction (qPCR) to determine if expression levels correlate with an increase in gonadal size. It will also be of interest to employ pharmacological techniques to investigate KP-type peptide function in *A. rubens*; for example, injection of KP-type peptides could be used to determine whether KP triggers gonadal growth and maturation in starfish.

The function of KP-type peptides (in addition to other novel neuropeptides) could also be investigated through the development of loss-of-function assays that allow the targeted knockout of genes in starfish. To date, although there is no genome available for the starfish species *A. rubens*, there is genome data available for the starfish species *P. miniata* (http://www.echinobase.org/Echinobase/pm). Nuclease-based genome-engineering technologies including zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered

regularly interspaced short palindromic repeats (CRISPR) have been utilised across a number of model systems including in the echinoderms. For example, ZFNs and TALENs have been used to investigate gene function in the embryo of the sea urchin *Hemicentrotus pulcherrimus* (Hosoi et al., 2014; Ochiai et al., 2010). However, with the sea urchin embryo a well-established tool in developmental biology, gene knockout approaches in the starfish *A. rubens* would have to be developed for the larval stage in order to allow delivery to the starfish germ line. The development of targeted gene knockouts will be crucial to furthering our understanding of neuropeptide function in non-model organisms, which do not possess the genetic toolbox available for classic model organisms including *Drosophila* (Mohr et al., 2014) and *C. elegans* (Boulin and Hobert, 2012).

7. Acknowledgements

I'd like to first and foremost thank my supervisor Prof. Maurice Elphick (School of Biological and Chemical Sciences (SBCS), QMUL) for his excellent guidance and support throughout my PhD studentship, which was funded by a Graduate Teaching Studentship (GTS) at SBCS, QMUL. I would also like to thank my co-supervisor Prof. Paul Chapple (William Harvey Research Institute (WHRI), QMUL) and panel members Dr. Rachel Ashworth (School of Medicine and Dentistry (SMD), QMUL) and Dr. Paola Oliveri (UCL) for their valuable advice during the course of my studentship. I would also like to acknowledge a number of colleagues who have contributed to each of the chapters of the thesis (in addition to the contribution of co-authors in publications arising out of the thesis).

7.1. Identification of transcripts encoding neuropeptide precursors and neuropeptide receptors in the starfish *A. rubens*

I'd like to thank the Barts and the London Genome Centre for sequencing the neural transcriptome of the starfish *Asterias rubens* (and the transcriptomes of the brittle star *Ophionotus victoriae* and the feather star *Antedon mediterranea* outlined in Elphick et al., 2015). In particular, I'd like to thank Dr. Monika Struebig (SBCS, QMUL) for her expertise in library preparation. I'd like to thank Mahesh Pancholi (SBCS, QMUL) for assembling the neural transcriptome of the starfish *A. rubens* and Liisa Blowes (SBCS, QMUL) for assembling the transcriptomes of the brittle star *O. victoriae* and the feather star *A. mediterranea*. Finally, I'd like to thank the contributions of Dr. Olivier Mirabeau (Cancer Genetics Unit, Institut Curie) and Ismail Moghul (former undergraduate student at SBCS, QMUL) for the identification of a number of novel neuropeptide precursors (that could not be

identified on the basis of BLAST homology) through a number of conversations and by using a hidden Markov model (HMM) and the prediction tool NPsearch (https://rubygems.org/gems/NpSearch) respectively.

7.2. Discovery and functional characterisation of a novel vasopressin/oxytocin (VP/OT)-type neuropeptide ("asterotocin") in the starfish *A. rubens*

Cloning of the asterotocin precursor was completed alongside former project student and current PhD student Esther Odekunle (SBCS, QMUL). Mass spectrometric confirmation of the asterotocin peptide was completed through collaboration with Dr. Susan E. Slade (University of Warwick). Both *in vitro* and *in vivo* functional studies were performed alongside former project student Natalia Martynuk (now a PhD student at the University of Cambridge). I'd also like to thank Ray Crundwell (SBCS, QMUL) for photographing and filming the asterotocin *in vivo* experiments and Dr. Michaela Egertová (SBCS, QMUL) for her valuable expertise and advice with regards to performing mRNA *in situ* hybridisation (ISH).

7.3. Discovery and functional characterisation of a novel neurophysin-associated neuropeptide ("NGFFYamide") in the starfish *A. rubens*

In vitro functional studies were performed by former project student Robyn Dane (now a student at Warwick Medical School, University of Warwick). Mass spectrometric confirmation of the NGFFYamide peptide was completed through collaboration with Prof. James. H. Scrivens and Dr. Susan E. Slade (University of Warwick). Once more, I'd also like to thank Ray Crundwell (SBCS, QMUL) for photographing and filming the NGFFYamide *in vivo* experiments and Dr. Michaela Egertová (SBCS, QMUL) for her valuable expertise and advice with regards to

performing mRNA ISH. Finally, I'd like to thank Dr. Matthew Parker (SBCS, QMUL) for his help in performing the statistical analysis for *in vitro* pharmacology experiments.

7.4. Discovery of the NGFFFamide receptor in the sea urchin *S. purpuratus* unites a bilaterian neuropeptide family

I'd like to thank Dr. Lilliane Schoofs (KU Leuven) for the opportunity to characterise the NGFFFamide receptor in her laboratory and Dr. Isabel Beets (KU Leuven) for her valuable expertise in performing the receptor deorphanisation assay and the subsequent data analysis. Finally, I'd like to thank the Company of Biologists (The *Journal of Experimental Biology*) for funding a Travelling Fellowship to complete the work in Dr. Lilliane Schoofs' laboratory.

8. References

Abdel-Latief, M. and Hoffmann, K. H. (2007). The adipokinetic hormones in the fall armyworm, *Spodoptera frugiperda*: cDNA cloning, quantitative real time RT-PCR analysis, and gene specific localization. *Insect Biochem Mol Biol* **37**, 999-1014.

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F. et al. (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185-2195.

Agüera, A., Trommelen, M., Burrows, F., Jansen, J. M., Schellekens, T. and Smaal, A. (2012). Winter feeding activity of the common starfish (*Asterias rubens* L.): the role of temperature and shading. *J Sea Res* **72**, 106-112.

Aikins, M. J., Schooley, D. A., Begum, K., Detheux, M., Beeman, R. W. and Park, Y. (2008). Vasopressin-like peptide and its receptor function in an indirect diuretic signaling pathway in the red flour beetle. *Insect Biochem Mol Biol* **38**, 740-748.

Aloyz, R. S. and DesGroseillers, L. (1995). Processing of the L5-67 precursor peptide and characterization of LUQIN in the LUQ neurons of *Aplysia californica*. *Peptides* **16**, 331-338.

Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. and Evans, R. M. (1982). Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature* **298**, 240-244.

Anderson, J. M. (1954). Studies on the cardiac stomach of the starfish, *Asterias forbesi*. *Biol Bull* **107**, 157-173.

Anderson, S. C., Flemming, J. M., Watson, R. and Lotze, H. K. (2011). Serial exploitation of global sea cucumber fisheries. *Fish and Fisheries* **12**, 317-339.

Arai, A. C. (2009). The role of kisspeptin and GPR54 in the hippocampus. *Peptides* **30**, 16-25.

Arakane, Y., Li, B., Muthukrishnan, S., Beeman, R. W., Kramer, K. J. and Park, Y. (2008). Functional analysis of four neuropeptides, EH, ETH, CCAP and bursicon, and their receptors in adult ecdysis behavior of the red flour beetle, *Tribolium castaneum*. *Mech Dev* **125**, 984-995.

Babey, M., Kopp, P. and Robertson, G. L. (2011). Familial forms of diabetes insipidus: clinical and molecular characteristics. *Nat Rev Endocrinol* **7**, 701-714.

Bahn, J. H., Lee, G. and Park, J. H. (2009). Comparative analysis of Pdf-mediated circadian behaviors between *Drosophila melanogaster* and *D. virilis*. *Genetics* **181**, 965-975.

Bai, H. and Palli, S. R. (2010). Functional characterization of bursicon receptor and genome-wide analysis for identification of genes affected by bursicon receptor RNAi. *Dev Biol* **344**, 248-258.

Baker, J. D. and Truman, J. W. (2002). Mutations in the *Drosophila* glycoprotein hormone receptor, rickets, eliminate neuropeptide-induced tanning and selectively block a stereotyped behavioral program. *J Exp Biol* **205**, 2555-2565.

Barker, M. F. and Nichols, D. (1983). Reproduction, recruitment and juvenile ecology of the starfish, *Asterias rubens* and *Marthasterias glacialis*. *J Mar Biol Assoc UK* **63**, 745-765.

Basch, P. F. (1956). Observations on the retractor strands of the starfish stomach. *Biol Rev City Coll NY* **18**, 14-17.

Bates, D. M. and Sarkar, D. (2007). lme4: Linear mixed-effects models using S4 classes, R package version. Available at: http://lme4.r-forge.r-project.org/.

Beck, B. (2006). Neuropeptide Y in normal eating and in genetic and dietary-induced obesity. *Philos Trans R Soc Lond B Biol Sci* **361**, 1159-1185.

Beets, I., Janssen, T., Meelkop, E., Temmerman, L., Suetens, N., Rademakers, S., Jansen, G. and Schoofs, L. (2012). Vasopressin/oxytocin-related signaling regulates gustatory associative learning in *C. elegans*. *Science* **338**, 543-545.

Beets, I., Temmerman, L., Janssen, T. and Schoofs, L. (2013). Ancient neuromodulation by vasopressin/oxytocin-related peptides. *Worm* **2**, e24246.

Beijnink, F. B. and Voogt, P. A. (1986). The aboral haemal system of the sea star, *Asterias rubens* (Echinodermata, Asteroidea): an ultrastructural and histochemical study. *Zoomorphology* **106**, 49-60.

Belmont, M., Cazzamali, G., Williamson, M., Hauser, F. and Grimmelikhuijzen, C. J. (2006). Identification of four evolutionarily related G protein-coupled receptors from the malaria mosquito *Anopheles gambiae*. *Biochem Biophys Res Commun* **344**, 160-165.

Bendena, W. G. (2010). Neuropeptide physiology in insects. *Adv Exp Med Biol* **692**, 166-191.

Bendtsen, J. D., Nielsen, H., von Heijne, G. and Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**, 783-795.

Bennett, H. P. J., Bradbury, W. B., Huttner, W. B. and Smyth, D. G. (1993). Processing of pro-peptides: glycosylation, phosphorylation, sulfation, acetylation and amidation. In: *Mechanisms of intracellular trafficking and processing of prohormones* (ed. **Loh, Y. P.**). pp. 251-288. Florida: CRC Press.

Birenheide, R., Tamori, M., Motokawa, T., Ohtani, M., Iwakoshi, E., Muneoka, Y., Fujita, T., Minakata, H. and Nomoto, K. (1998). Peptides controlling stifness of connective tissue in sea cucumbers. *Biol Bull* **194**, 253-259.

Birnbaumer, M., Seibold, A., Gilbert, S., Ishido, M., Barberis, C., Antaramian, A., Brabet, P. and Rosenthal, W. (1992). Molecular cloning of the receptor for human antidiuretic hormone. *Nature* **357**, 333-335.

Bockaert, J. and Pin, J. P. (1999). Molecular tinkering of G protein-coupled receptors: an evolutionary success. *Embo J* **18**, 1723-1729.

Boler, J., Enzmann, F., Folkers, K., Bowers, C. Y. and Schally, A. V. (1969). The identity of chemical and hormonal properties of the thyrotropin releasing hormone and pyroglutamyl-histidyl-proline amide. *Biochem Biophys Res Commun* **37**, 705-710.

Bondy, C. A., Gainer, H. and Russell, J. T. (1987). Effects of stimulus frequency and potassium channel blockade on the secretion of vasopressin and oxytocin from the neurohypophysis. *Neuroendocrinology* **46**, 258-267.

Boulin, T. and Hobert, O. (2012). From genes to function: the *C. elegans* genetic toolbox. *Wiley Interdiscip Rev Dev Biol* **1**, 114-137.

Bourlat, S. J., Juliusdottir, T., Lowe, C. J., Freeman, R., Aronowicz, J., Kirschner, M., Lander, E. S., Thorndyke, M., Nakano, H., Kohn, A. B. et al. (2006). Deuterostome phylogeny reveals monophyletic chordates and the new phylum *Xenoturbellida*. *Nature* **444**, 85-88.

Brazeau, P., Vale, W., Burgus, R., Ling, N., Butcher, M., Rivier, J. and Guillemin, R. (1973). Hypothalamic polypeptide that inhibits the secretion of immunoreactive pituitary growth hormone. *Science* **179**, 77-79.

Breese, G. R., Mueller, R. A., Mailman, R. B. and Frye, G. D. (1981). Effects of TRH on central nervous system function. *Prog Clin Biol Res* **68**, 99-116.

Broad, K. D., Curley, J. P. and Keverne, E. B. (2006). Mother-infant bonding and the evolution of mammalian social relationships. *Philos Trans R Soc Lond B Biol Sci* **361**, 2199-2214.

Brody, T. and Cravchik, A. (2000). *Drosophila melanogaster* G protein-coupled receptors. *J Cell Biol* **150**, F83-88.

Broertjes, J. J. S. and Posthuma, G. (1978). Direct visualization of the haemal system in starfish by a staining procedure. *Experientia* **34**, 1243-1245.

Bromham, L. D. and Degnan, B. M. (1999). Hemichordates and deuterostome evolution: robust molecular phylogenetic support for a hemichordate + echinoderm clade. *Evol Dev* **1**, 166-171.

Brownstein, M. J., Russell, J. T. and Gainer, H. (1980). Synthesis, transport, and release of posterior pituitary hormones. *Science* **207**, 373-378.

Brusca, R. C. and Brusca, G. J. (2003). Invertebrates, 2nd edition. Sunderland, MA: Sinauer Associates.

Bullesbach, E. E. and Schwabe, C. (2000). The relaxin receptor-binding site geometry suggests a novel gripping mode of interaction. *J Biol Chem* **275**, 35276-35280.

Burbach, J. P. (2012). What are neuropeptides? In: *Neuropeptides: methods and protocols, vol 789*. (ed. **Merighi., A.**) pp. 1-36. New York: Humana Press.

Burke, R. D., Angerer, L. M., Elphick, M. R., Humphrey, G. W., Yaguchi, S., Kiyama, T., Liang, S., Mu, X., Agca, C., Klein, W. H. et al. (2006). A genomic view of the sea urchin nervous system. *Dev Biol* **300**, 434-460.

Burnett, A. L. and Anderson, J. M. (1955). The contractile properties of the retractor mechanism of the cardiac stomach in *Asterias forbesi*. *Anat Rec* **122**, 463–464.

Byrne, M. (2001). The morphology of autotomy structures in the sea cucumber *Eupentacta quinquesemita* before and during evisceration. *J Exp Biol* **204**, 849-863.

Cabrero, P., Radford, J. C., Broderick, K. E., Costes, L., Veenstra, J. A., Spana, E. P., Davies, S. A. and Dow, J. A. (2002). The Dh gene of *Drosophila melanogaster* encodes a diuretic peptide that acts through cyclic AMP. *J Exp Biol* **205**, 3799-3807.

Caldwell, H. K., Lee, H. J., Macbeth, A. H. and Young, W. S., 3rd. (2008). Vasopressin: behavioral roles of an "original" neuropeptide. *Prog Neurobiol* **84**, 1-24.

Campbell, R. K., Satoh, N. and Degnan, B. M. (2004). Piecing together evolution of the vertebrate endocrine system. *Trends Genet* **20**, 359-366.

Carter, C. S. (2014). Oxytocin pathways and the evolution of human behavior. *Annu Rev Psychol* **65**, 17-39.

Carter, C. S., DeVries, A. C. and Getz, L. L. (1995). Physiological substrates of mammalian monogamy: the prairie vole model. *Neurosci Biobehav Rev* **19**, 303-314.

Cazzamali, G., Hauser, F., Kobberup, S., Williamson, M. and Grimmelikhuijzen, C. J. (2003). Molecular identification of a *Drosophila* G protein-coupled receptor specific for crustacean cardioactive peptide. *Biochem Biophys Res Commun* **303**, 146-152.

Chaet, A. B. (1964). A mechanism for obtaining mature gametes from starfish. *Biol Bull* **126**, 8-13.

Chaet, A. B., McConnaughy R. A. (1959). Physiologic activity of nerve extracts. *Biol Bull* **117**, 407–408.

Champagne, D. E. and Ribeiro, J. M. (1994). Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito *Aedes aegypti*. *Proc Natl Acad Sci U S A* **91**, 138-142.

Chang, M. M. and Leeman, S. E. (1970). Isolation of a sialogogic peptide from bovine hypothalamic tissue and its characterization as substance P. *J Biol Chem* **245**, 4784-4790.

Chen, X., Peterson, J., Nachman, R. J. and Ganetzky, B. (2012). Drosulfakinin activates CCKLR-17D1 and promotes larval locomotion and escape response in *Drosophila*. *Fly (Austin)* **6**, 290-297.

Cho, M. M., DeVries, A. C., Williams, J. R. and Carter, C. S. (1999). The effects of oxytocin and vasopressin on partner preferences in male and female prairie voles (*Microtus ochrogaster*). *Behav Neurosci* **113**, 1071-1079.

Christie, A. E., Lundquist, C. T., Nässel, D. R. and Nusbaum, M. P. (1997). Two novel tachykinin-related peptides from the nervous system of the crab *Cancer borealis*. *J Exp Biol* **200**, 2279-2294.

Chung, J. S., Katayama, H. and Dircksen, H. (2012). New functions of arthropod bursicon: inducing deposition and thickening of new cuticle and hemocyte granulation in the blue crab, *Callinectes sapidus*. *PLoS One* **7**, e46299.

Chung, J. S. and Webster, S. G. (2004). Expression and release patterns of neuropeptides during embryonic development and hatching of the green shore crab, *Carcinus maenas*. *Development* **131**, 4751-4761.

Cobb, J. L. (1970). The significance of the radial nerve cords in asteroids and echinoids. *Z Zellforsch Mikrosk Anat* **108**, 457-474.

Cobb, J. L. (1978). An ultrastructural study of the dermal papulae of the starfish, *Asterias rubens*, with special reference to innervation of the muscles. *Cell Tissue Res* **187**, 515-523.

Cobb, J. L. S. (1987). Neurobiology of Echinodermata. In: *Nervous systems of invertebrates* (ed. **Ali, M. A.**). pp. 483-525. New York: Plenum Press.

Colbourne, J. K., Pfrender, M. E., Gilbert, D., Thomas, W. K., Tucker, A., Oakley, T. H., Tokishita, S., Aerts, A., Arnold, G. J., Basu, M. K. et al. (2011). The ecoresponsive genome of *Daphnia pulex*. *Science* **331**, 555-561.

Colledge, W. H. (2009). Transgenic mouse models to study Gpr54/kisspeptin physiology. *Peptides* **30**, 34-41.

Conn, P. J. and Kaczmarek, L. K. (1989). The bag cell neurons of *Aplysia*. A model for the study of the molecular mechanisms involved in the control of prolonged animal behaviors. *Mol Neurobiol* **3**, 237-273.

Conzelmann, M., Offenburger, S. L., Asadulina, A., Keller, T., Munch, T. A. and Jekely, G. (2011). Neuropeptides regulate swimming depth of *Platynereis* larvae. *Proc Natl Acad Sci U S A* **108**, e1174-83.

Cool, D. R., Jackson, S. B. and Waddell, K. S. (2008). Structural requirements for sorting pro-vasopressin to the regulated secretory pathway in a neuronal cell line. *Open Neuroendocrinol J* **1**, 1-8.

Cruz, L. J., de Santos, V., Zafaralla, G. C., Ramilo, C. A., Zeikus, R., Gray, W. R. and Olivera, B. M. (1987). Invertebrate vasopressin/oxytocin homologs. Characterization of peptides from *Conus geographus* and *Conus straitus* venoms. *J Biol Chem* **262**, 15821-15824.

Cuénot, L. (1948). Anatomie, Ethologie et Systématique des Echinodermes. In: *Traité de Zoologie, vol. 11* (ed. **Grassé, P. P.**). pp. 1-363. Paris: Masson.

Cuénot, L. (1886). Sur les functions de la glande ovoïde, des corps de Tiedemann et des vésicules de Poli chez les Astérides. *Comptes rendus de l'Académie des Sciences*. **102,** 1568–1569.

d'Anglemont de Tassigny, X., Fagg, L. A., Dixon, J. P., Day, K., Leitch, H. G., Hendrick, A. G., Zahn, D., Franceschini, I., Caraty, A., Carlton, M. B. et al. (2007). Hypogonadotropic hypogonadism in mice lacking a functional Kiss1 gene. *Proc Natl Acad Sci U S A* **104**, 10714-10719.

Dare, P. J. (1982). Notes on the swarming behaviour and population density of *Asterias rubens* L. (Echinodermata: Asteroidea) feeding on the mussel, *Mytilus edulis* L. *Journal du Conseil* **40**, 112-118.

de Bree, F. M. (2000). Trafficking of the vasopressin and oxytocin prohormone through the regulated secretory pathway. *J Neuroendocrinol* **12**, 589-594.

de Bree, F. M. and Burbach, J. P. (1998). Structure-function relationships of the vasopressin prohormone domains. *Cell Mol Neurobiol* **18**, 173-191.

de Lecea, L., del Rio, J. A., Criado, J. R., Alcantara, S., Morales, M., Danielson, P. E., Henriksen, S. J., Soriano, E. and Sutcliffe, J. G. (1997). Cortistatin is expressed in a distinct subset of cortical interneurons. *J Neurosci* **17**, 5868-5880.

de Lecea, L., Kilduff, T. S., Peyron, C., Gao, X., Foye, P. E., Danielson, P. E., Fukuhara, C., Battenberg, E. L., Gautvik, V. T., Bartlett, F. S., 2nd et al. (1998). The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. *Proc Natl Acad Sci U S A* **95**, 322-327.

de Pablo, F., Chambers, S. A. and Ota, A. (1988). Insulin-related molecules and insulin effects in the sea urchin embryo. *Dev Biol* **130**, 304-310.

De'ath, G., Fabricius, K. E., Sweatman, H. and Puotinen, M. (2012). The 27-year decline of coral cover on the Great Barrier Reef and its causes. *Proc Natl Acad Sci U S A* **109**, 17995-17999.

Dehal, P., Satou, Y., Campbell, R. K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D. M. et al. (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science* **298**, 2157-2167.

Deng, C., Reddy, P., Cheng, Y., Luo, C. W., Hsiao, C. L. and Hsueh, A. J. (2013). Multi-functional norrin is a ligand for the LGR4 receptor. *J Cell Sci* **126**, 2060-2068.

Devi, L. (1991). Consensus sequence for processing of peptide precursors at monobasic sites. *FEBS Lett* **280**, 189-194.

Di Cristo, C., De Lisa, E. and Di Cosmo, A. (2009). Control of GnRH expression in the olfactory lobe of *Octopus vulgaris*. *Peptides* **30**, 538-544.

Dockray, G. J. (1977). Molecular evolution of gut hormones: application of comparative studies on the regulation of digestion. *Gastroenterology* **72**, 344-358.

Dockray, G. J, Gregory, R. A., Hutchison, J. B., Harris, J. I. and Runswick, M. J. (1978) Isolation, structure and biological activity of two cholecystokinin octapeptides from sheep brain. *Nature*, **274**, 711-713.

Dolmer, P. (1998). The interactions between bed structure of *Mytilus edulis* L. and the predator *Asterias rubens* L. *Journal of Experimental Marine Biology and Ecology* **228**, 137-150.

Domschke, K., Reif, A., Weber, H., Richter, J., Hohoff, C., Ohrmann, P., Pedersen, A., Bauer, J., Suslow, T., Kugel, H. et al. (2011). Neuropeptide S receptor gene - converging evidence for a role in panic disorder. *Mol Psychiatry* **16**, 938-948.

Donaldson, Z. R. and Young, L. J. (2008). Oxytocin, vasopressin, and the neurogenetics of sociality. *Science* **322**, 900-904.

Dores, R. M., Steveson, T. C. and Lopez, K. (1991). Differential mechanisms for the N-acetylation of alpha-melanocyte-stimulating hormone and beta-endorphin in the intermediate pituitary of the frog, *Xenopus laevis*. *Neuroendocrinology* **53**, 54-62.

Dos Santos, S., Bardet, C., Bertrand, S., Escriva, H., Habert, D. and Querat, B. (2009). Distinct expression patterns of glycoprotein hormone-alpha2 and -beta5 in a basal chordate suggest independent developmental functions. *Endocrinology* **150**, 3815-3822.

Downes, G. B. and Gautam, N. (1999). The G protein subunit gene families. *Genomics* **62**, 544-552.

Dreifuss, J. J., Kalnins, I., Kelly, J. S. and Ruf, K. B. (1971). Action potentials and release of neurohypophysial hormones *in vitro*. *J Physiol* **215**, 805-17.

Drews, J. (2000). Drug discovery: a historical perspective. *Science* **287**, 1960-1964.

Du, H., Bao, Z., Hou, R., Wang, S., Su, H., Yan, J., Tian, M., Li, Y., Wei, W., Lu, W. et al. (2012). Transcriptome sequencing and characterization for the sea cucumber *Apostichopus japonicus* (Selenka, 1867). *PLoS One* **7**, e33311.

Dufresne, M., Seva, C. and Fourmy, D. (2006). Cholecystokinin and gastrin receptors. *Physiol Rev* **86**, 805-847.

Eipper, B. A., Milgram, S. L., Husten, E. J., Yun, H. Y. and Mains, R. E. (1993). Peptidylglycine alpha-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains. *Protein Sci* **2**, 489-497.

Elekonich, M. M. and Horodyski, F. M. (2003). Insect allatotropins belong to a family of structurally-related myoactive peptides present in several invertebrate phyla. *Peptides* **24**, 1623-1632.

Elizur, A. (2009). The KiSS1/GPR54 system in fish. *Peptides* **30**, 164-70.

Elphick, M. R. (2010). NG peptides: a novel family of neurophysin-associated neuropeptides. *Gene* **458**, 20-26.

Elphick, M. R. (2012). The protein precursors of peptides that affect the mechanics of connective tissue and/or muscle in the echinoderm *Apostichopus japonicus*. *PLoS One* **7**, e44492.

Elphick, M. R. (2014). SALMFamide salmagundi: the biology of a neuropeptide family in echinoderms. *Gen Comp Endocrinol* **205**, 23-35.

Elphick, M. R., Achhala, S. and Martynyuk, N. (2013). The evolution and diversity of SALMFamide neuropeptides. *PLoS One* **8**, e59076.

Elphick, M. R. and Melarange, R. (1998). Nitric oxide function in an echinoderm. *Biological Bulletin* **194**, 260-266.

Elphick, M. R. and Melarange, R. (2001). Neural control of muscle relaxation in echinoderms. *J Exp Biol* **204**, 875-885.

Elphick, M. R., Newman, S. J. and Thorndyke, M. C. (1995). Distribution and action of SALMFamide neuropeptides in the starfish *Asterias rubens*. *J Exp Biol* **198**, 2519-2525.

Elphick, M. R., Price, D. A., Lee, T. D. and Thorndyke, M. C. (1991a). The SALMFamides: a new family of neuropeptides isolated from an echinoderm. *Proc Biol Sci* **243**, 121-127.

Elphick, M. R., Reeve, J. R., Jr., Burke, R. D. and Thorndyke, M. C. (1991b). Isolation of the neuropeptide SALMFamide-1 from starfish using a new antiserum. *Peptides* **12**, 455-459.

Elphick, M. R., Roland, H. E. and Thorndyke, M. C. (1989). FMRFamide-like immunoreactivity in the nervous system of the starfish *Asterias rubens*. *Biological Bulletin* **177**, 141-145.

Elphick, M. R. and Rowe, M. L. (2009). NGFFFamide and echinotocin: structurally unrelated myoactive neuropeptides derived from neurophysin-containing precursors in sea urchins. *J Exp Biol* **212**, 1067-1077.

Elphick, M. R., Semmens, D. C., Blowes, L. M., Levine, J., Lowe, C. J., Arnone, M. I. and Clark, M. S. (2015). Reconstructing SALMFamide neuropeptide precursor evolution in the phylum Echinodermata: ophiuroid and crinoid sequence data provide new insights. *Front Endocrinol (Lausanne)* **6**, 2.

Elphick, M. R. and Thorndyke, M. C. (2005). Molecular characterisation of SALMFamide neuropeptides in sea urchins. *J Exp Biol* **208**, 4273-4282.

Epelbaum, J. (1986). Somatostatin in the central nervous system: physiology and pathological modifications. *Prog Neurobiol* **27**, 63-100.

Eugenin, E. A., Gonzalez, H., Saez, C. G. and Saez, J. C. (1998). Gap junctional communication coordinates vasopressin-induced glycogenolysis in rat hepatocytes. *Am J Physiol* **274**, G1109-1116.

Ewer, J., Wang, C. M., Klukas, K. A., Mesce, K. A., Truman, J. W. and Fahrbach, S. E. (1998). Programmed cell death of identified peptidergic neurons involved in ecdysis behavior in the Moth, *Manduca sexta*. *J Neurobiol* **37**, 265-280.

Fiddes, J. C. and Goodman, H. M. (1979). Isolation, cloning and sequence analysis of the cDNA for the alpha-subunit of human chorionic gonadotropin. *Nature* **281**, 351-356.

Fiddes, J. C. and Goodman, H. M. (1980). The cDNA for the beta-subunit of human chorionic gonadotropin suggests evolution of a gene by readthrough into the 3'-untranslated region. *Nature* **286**, 684-687.

Flammang, P. (1996). Adhesion in echinoderms. In: *Echinoderm Studies, vol. 5* (ed. **Jangoux, M. and Lawrence, J. M.**). pp. 1–60. Rotterdam: Balkema.

Fodor, M., van Elk, E. J., Huizinga, C. T., Prins, T. and Delemarre-van de Waal, H. A. (2005). NPY neurons express somatostatin receptor subtype 1 in the arcuate nucleus. *Neuroreport* **16**, 29-32.

Fraenkel, G., Hsiao, C. and Seligman, M. (1966). Properties of bursicon: an insect protein hormone that controls cuticular tanning. *Science* **151**, 91-93.

Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. and Schioth, H. B. (2003). The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol* **63**, 1256-1272.

Freeman, R. M., Jr., Wu, M., Cordonnier-Pratt, M. M., Pratt, L. H., Gruber, C. E., Smith, M., Lander, E. S., Stange-Thomann, N., Lowe, C. J., Gerhart, J. et al. (2008). cDNA sequences for transcription factors and signaling proteins of the hemichordate *Saccoglossus kowalevskii*: efficacy of the expressed sequence tag (EST) approach for evolutionary and developmental studies of a new organism. *Biol Bull* **214**, 284-302.

Fujimoto, K., Ohta, N., Yoshida, M., Kubota, I., Muneoka, Y. and Kobayashi, M. (1990). A novel cardio-excitatory peptide isolated from the atria of the African giant snail, *Achatina fulica*. *Biochem Biophys Res Commun* **167**, 777-783.

Fujino, Y., Nagahama, T., Oumi, T., Ukena, K., Morishita, F., Furukawa, Y., Matsushima, O., Ando, M., Takahama, H., Satake, H. et al. (1999). Possible functions of oxytocin/vasopressin-superfamily peptides in annelids with special reference to reproduction and osmoregulation. *J Exp Zool* **284**, 401-406.

Fujisawa, J., Muneaka, Y., Tabahashi, T., Takao, T., Shimonishi, Y., Kubota, I., Ikeda, T., Minakata, H., Nomoto, K., Kiss, T. et al. (1993). An invertebrate-type tachykinin isolated from the freshwater bivalve mollusc, *Anodonta cygnea*. In: *Peptide Chemistry* (ed. **Okoda, Y.**). pp. 161–164. Osaka: Protein Research Foundation.

Funes, S., Hedrick, J. A., Vassileva, G., Markowitz, L., Abbondanzo, S., Golovko, A., Yang, S., Monsma, F. J. and Gustafson, E. L. (2003). The KiSS-1 receptor GPR54 is essential for the development of the murine reproductive system. *Biochem Biophys Res Commun* **312**, 1357-1363.

Furuya, K., Milchak, R. J., Schegg, K. M., Zhang, J., Tobe, S. S., Coast, G. M. and Schooley, D. A. (2000). Cockroach diuretic hormones: characterization of a calcitonin-like peptide in insects. *Proc Natl Acad Sci U S A* **97**, 6469-6474.

Gade, G. (2004). Regulation of intermediary metabolism and water balance of insects by neuropeptides. *Annu Rev Entomol* **49**, 93-113.

Gainer, H., Wolfe, S. A., Jr., Obaid, A. L. and Salzberg, B. M. (1986). Action potentials and frequency-dependent secretion in the mouse neurohypophysis. *Neuroendocrinology* **43**, 557-563.

Galas, L., Raoult, E., Tonon, M. C., Okada, R., Jenks, B. G., Castano, J. P., Kikuyama, S., Malagon, M., Roubos, E. W. and Vaudry, H. (2009). TRH acts as a multifunctional hypophysiotropic factor in vertebrates. *Gen Comp Endocrinol* **164**, 40-50.

Gammie, S. C. and Truman, J. W. (1997). Neuropeptide hierarchies and the activation of sequential motor behaviors in the hawkmoth, *Manduca sexta*. *J Neurosci* **17**, 4389-4397.

Garrison, J. L., Macosko, E. Z., Bernstein, S., Pokala, N., Albrecht, D. R. and Bargmann, C. I. (2012). Oxytocin/vasopressin-related peptides have an ancient role in reproductive behavior. *Science* **338**, 540-543.

Geraerts, W. P. (1992). Neurohormonal control of growth and carbohydrate metabolism by the light green cells in *Lymnaea stagnalis*. *Gen Comp Endocrinol* **86**, 433-444.

Gimpl, G. and Fahrenholz, F. (2001). The oxytocin receptor system: structure, function, and regulation. *Physiol Rev* **81**, 629-683.

Gonnet, G. H., Cohen, M. A. and Benner, S. A. (1992). Exhaustive matching of the entire protein sequence database. *Science* **256**, 1443-1435.

Good-Avila, S. V., Yegorov, S., Harron, S., Bogerd, J., Glen, P., Ozon, J. and Wilson, B. C. (2009). Relaxin gene family in teleosts: phylogeny, syntenic mapping, selective constraint, and expression analysis. *BMC Evol Biol* **9**, 293.

Gorbman, A. and Sower, S. A. (2003). Evolution of the role of GnRH in animal (Metazoan) biology. *Gen Comp Endocrinol* **134**, 207-213.

Grimmelikhuijzen, C. P., Williamson, M. and Hansen, G. (2004). Neuropeptides in cnidarians. In: *Cell Signalling in Prokaryotes and Lower Metazoa*, (ed. **Fairweather, I**.). pp. 115-139. Netherlands: Springer.

Guilgur, L. G., Moncaut, N. P., Canario, A. V. and Somoza, G. M. (2006). Evolution of GnRH ligands and receptors in gnathostomata. *Comp Biochem Physiol A Mol Integr Physiol* **144**, 272-283.

Guilloteau, P., Le Meuth-Metzinger, V., Morisset, J. and Zabielski, R. (2006). Gastrin, cholecystokinin and gastrointestinal tract functions in mammals. *Nutr Res Rev* **19**, 254-283.

Gupte, J., Cutler, G., Chen, J. L. and Tian, H. (2004). Elucidation of signaling properties of vasopressin receptor-related receptor 1 by using the chimeric receptor approach. *Proc Natl Acad Sci U S A* **101**, 1508-1513.

Gwee, P. C., Tay, B. H., Brenner, S. and Venkatesh, B. (2009). Characterization of the neurohypophysial hormone gene loci in elephant shark and the Japanese lamprey: origin of the vertebrate neurohypophysial hormone genes. *BMC Evol Biol* **9**, 47.

Hall, J. D. and Lloyd, P. E. (1990). Involvement of pedal peptide in locomotion in *Aplysia*: modulation of foot muscle contractions. *J Neurobiol* **21**, 858-868.

Harada, A., Yoshida, M., Minakata, H., Nomoto, K., Muneoka, Y. and Kobayashi, M. (1993). Structure and function of the molluscan myoactive tetradecapeptides. *Zoolog Sci* **10**, 257-265.

Harms, J. F., Welch, D. R. and Miele, M. E. (2003). KISS1 metastasis suppression and emergent pathways. *Clin Exp Metastasis* **20**, 11-18.

Harshini, S., Nachman, R. J. and Sreekumar, S. (2002). Inhibition of digestive enzyme release by neuropeptides in larvae of *Opisina arenosella* (Lepidoptera: Cryptophasidae). *Comp Biochem Physiol B Biochem Mol Biol* **132**, 353-358.

Hayashi, R. (1935). Studies on the morphology of Japanese sea-stars. II. Internal anatomy of two short-rayed sea-stars, *Patiria pectinifera* (Müller and Troschel) and *Asterina batheri* (Goto). *J Fac Sci Hokkaido Imp Univ Ser VI, Zool.* **4**, 197–212.

Henderson, K. K. and Byron, K. L. (2007). Vasopressin-induced vasoconstriction: two concentration-dependent signaling pathways. *J Appl Physiol (1985)* **102**, 1402-1409.

Hennebert, E. (2010). Adhesion mechanisms developed by sea stars: a review of the ultrastructure and composition of tube feet and their secretion. In: *Biological Adhesive Systems* (ed. **von Byern, J. and Grunwald, I.**). pp. 99-109. Vienna: Springer.

Hewes, R. S. and Taghert, P. H. (2001). Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. *Genome Res* **11**, 1126-1142.

Hofer, S. and Homberg, U. (2006). Evidence for a role of orcokinin-related peptides in the circadian clock controlling locomotor activity of the cockroach *Leucophaea maderae*. *J Exp Biol* **209**, 2794-2803.

Hoffmann, F. G. and Opazo, J. C. (2011). Evolution of the relaxin/insulin-like gene family in placental mammals: implications for its early evolution. *J Mol Evol* **72**, 72-79.

Holland, L. Z., Albalat, R., Azumi, K., Benito-Gutierrez, E., Blow, M. J., Bronner-Fraser, M., Brunet, F., Butts, T., Candiani, S., Dishaw, L. J. et al. (2008). The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res* **18**, 1100-1111.

Holm, K., Dupont, S., Skold, H., Stenius, A., Thorndyke, M. and Hernroth, B. (2008). Induced cell proliferation in putative haematopoietic tissues of the sea star, *Asterias rubens* (L.). *J Exp Biol* **211**, 2551-2558.

Hopkins, A. L. and Groom, C. R. (2002). The druggable genome. *Nat Rev Drug Discov* **1**, 727-730.

Hordijk, P. L., Schallig, H. D., Ebberink, R. H., de Jong-Brink, M. and Joosse, J. (1991). Primary structure and origin of schistosomin, an anti-gonadotropic neuropeptide of the pond snail *Lymnaea stagnalis*. *Biochem J* **279**, 837-842.

Hosoi, S., Sakuma, T., Sakamoto, N. and Yamamoto, T. (2014). Targeted mutagenesis in sea urchin embryos using TALENs. *Dev Growth Differ* **56**, 92-97.

Howard, A. D., McAllister, G., Feighner, S. D., Liu, Q., Nargund, R. P., Van der Ploeg, L. H. and Patchett, A. A. (2001). Orphan G-protein-coupled receptors and natural ligand discovery. *Trends Pharmacol Sci* **22**, 132-140.

Hoyle, C. H. (1999). Neuropeptide families and their receptors: evolutionary perspectives. *Brain Res* **848**, 1-25.

Hruby, V. J., al-Obeidi, F. and Kazmierski, W. (1990). Emerging approaches in the molecular design of receptor-selective peptide ligands: conformational, topographical and dynamic considerations. *Biochem J* **268**, 249-262.

Hsu, S. Y., Nakabayashi, K. and Bhalla, A. (2002). Evolution of glycoprotein hormone subunit genes in bilateral metazoa: identification of two novel human glycoprotein hormone subunit family genes, GPA2 and GPB5. *Mol Endocrinol* **16**, 1538-1551.

Hyman, L. H. (1955). The invertebrates: Echinodermata, vol. 4. New York: McGraw-Hill.

Ida, T., Takahashi, T., Tominaga, H., Sato, T., Kume, K., Yoshizawa-Kumagaye, K., Nishio, H., Kato, J., Murakami, N., Miyazato, M. et al. (2011). Identification of the endogenous cysteine-rich peptide trissin, a ligand for an orphan G protein-coupled receptor in *Drosophila*. *Biochem Biophys Res Commun* **414**, 44-48.

Ikeda, T., Minakata, H., Nomoto, K., Kubota, I. and Muneoka, Y. (1993). Two novel tachykinin-related neuropeptides in the echiuroid worm, *Urechis unicinctus*. *Biochem Biophys Res Commun* **192**, 1-6.

Inoue, M., Birenheide, R., Koizumi, O., Kobayakawa, Y., Muneoka, Y. and Motokawa, T. (1999). Localization of the neuropeptide NGIWYamide in the holothurian nervous system and its effects on muscular contraction. *Proc R Soc B* **266**, 993-993.

Insel, T. R. and Young, L. J. (2001). The neurobiology of attachment. *Nat Rev Neurosci* **2**, 129-136.

Iwakoshi, E., Ohtani, M., Takahashi, T., Muneoka, Y., Ikeda, T., Fujita, T., Minakata, H. and Nomoto, K. (1995). Comparative aspects of structure and action of bioactive peptides isolated from the sea cucumber *Stichopus japonicus*. In: *Peptide Chemistry* (ed. **Ohno, M.**). pp. 261-264. Osaka: Protein Research Foundation.

Iwakoshi, E., Takuwa-Kuroda, K., Fujisawa, Y., Hisada, M., Ukena, K., Tsutsui, K. and Minakata, H. (2002). Isolation and characterization of a GnRH-like peptide from *Octopus vulgaris*. *Biochem Biophys Res Commun* **291**, 1187-1193.

Jackson, M. B., Konnerth, A. and Augustine, G. J. (1991). Action potential broadening and frequency-dependent facilitation of calcium signals in pituitary nerve terminals. *Proc Natl Acad Sci U S A* **88**, 380-384.

Janssen, T., Husson, S. J., Meelkop, E., Temmerman, L., Lindemans, M., Verstraelen, K., Rademakers, S., Mertens, I., Nitabach, M., Jansen, G. et al. (2009). Discovery and characterization of a conserved pigment dispersing factor-like neuropeptide pathway in *Caenorhabditis elegans*. *J Neurochem* **111**, 228-241.

Janssen, T., Meelkop, E., Lindemans, M., Verstraelen, K., Husson, S. J., Temmerman, L., Nachman, R. J. and Schoofs, L. (2008). Discovery of a cholecystokinin-gastrin-like signaling system in nematodes. *Endocrinology* **149**, 2826-2839.

Jekely, G. (2013). Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci U S A* **110**, 8702-8707.

Johnsen, A. H. (1998). Phylogeny of the cholecystokinin/gastrin family. *Front Neuroendocrinol* **19**, 73-99.

Johnsen, A. H. and Rehfeld, J. F. (1990). Cionin: a disulfotyrosyl hybrid of cholecystokinin and gastrin from the neural ganglion of the protochordate *Ciona intestinalis*. *J Biol Chem* **265**, 3054-3058.

Jones C. E., Zandawala M., Semmens D. C., Anderson S., Hanson G. R., Janies D. and Elphick M. R. (2016). Identification of a neuropeptide precursor protein that gives rise to a "cocktail" of peptides that bind Cu(II) and generate metal-linked dimers. *Biochim Biophys Acta* **1860**, 57-66.

Kah, O., Lethimonier, C., Somoza, G., Guilgur, L. G., Vaillant, C. and Lareyre, J. J. (2007). GnRH and GnRH receptors in metazoa: a historical, comparative, and evolutive perspective. *Gen Comp Endocrinol* **153**, 346-364.

Kamesh, N., Aradhyam, G. K. and Manoj, N. (2008). The repertoire of G protein-coupled receptors in the sea squirt *Ciona intestinalis*. *BMC Evol Biol* **8**, 129.

Kanatani, H., Ikegami, S., Shirai, H., Oide, H. and Tamura, S. (1971). Purification of gonad-stimulating substance obtained from radial nerves of the starfish, *Asterias amurensis*. *Dev Growth Differ* **13**, 151-164.

Kanatani, H. and Shirai, H. (1970). Mechanism of starfish spawning. III. Properties and action of meiosis-inducing substance produced in gonad under influence of gonad-stimulating substance. *Dev Growth Differ* **12**, 119-140.

Kanatani, H. (1985). Oocyte growth and maturation in starfish. In: *Biology of Fertilization, vol. 1* (ed. **Metz C .B. and Monroy, A.**). pp. 119-140. New York: Academic Press.

Kanda, A., Iwakoshi-Ukena, E., Takuwa-Kuroda, K. and Minakata, H. (2003). Isolation and characterization of novel tachykinins from the posterior salivary gland of the common octopus *Octopus vulgaris*. *Peptides* **24**, 35-43.

Kangawa, K., Minamino, N., Fukuda, A. and Matsuo, H. (1983). Neuromedin K: a novel mammalian tachykinin identified in porcine spinal cord. *Biochem Biophys Res Commun* **114**, 533-540.

Kataoka, H., Toschi, A., Li, J. P., Carney, R. L., Schooley, D. A. and Kramer, S. J. (1989). Identification of an allatotropin from adult *Manduca sexta*. *Science* **243**, 1481-1483.

Kato, S., Tsurumaru, S., Taga, M., Yamane, T., Shibata, Y., Ohno, K., Fujiwara, A., Yamano, K. and Yoshikuni, M. (2009). Neuronal peptides induce oocyte maturation and gamete spawning of sea cucumber, *Apostichopus japonicus*. *Dev Biol* **326**, 169-176.

Kaufmann, C. and Brown, M. R. (2006). Adipokinetic hormones in the African malaria mosquito, *Anopheles gambiae*: identification and expression of genes for two peptides and a putative receptor. *Insect Biochem Mol Biol* **36**, 466-481.

Kavanaugh, S. I., Nozaki, M. and Sower, S. A. (2008). Origins of gonadotropin-releasing hormone (GnRH) in vertebrates: identification of a novel GnRH in a basal vertebrate, the sea lamprey. *Endocrinology* **149**, 3860-3869.

Kawada, T., Satake, H., Minakata, H., Muneoka, Y. and Nomoto, K. (1999). Characterization of a novel cDNA sequence encoding invertebrate tachykinin-related peptides isolated from the echiuroid worm, *Urechis unicinctus*. *Biochem Biophys Res Commun* **263**, 848-852.

Kawada, T., Sekiguchi, T., Itoh, Y., Ogasawara, M. and Satake, H. (2008). Characterization of a novel vasopressin/oxytocin superfamily peptide and its receptor from an ascidian, *Ciona intestinalis*. *Peptides* **29**, 1672-1678.

Kawauchi, H. (2006). Functions of melanin-concentrating hormone in fish. *J Exp Zool A Comp Exp Biol* **305**, 751-760.

Kawauchi, H. and Baker, B. I. (2004). Melanin-concentrating hormone signaling systems in fish. *Peptides* **25**, 1577-1584.

Kawauchi, H., Kawazoe, I., Tsubokawa, M., Kishida, M. and Baker, B. I. (1983). Characterization of melanin-concentrating hormone in chum salmon pituitaries. *Nature* **305**, 321-323.

Kawazoe, I., Kawauchi, H., Hirano, T. and Naito, N. (1987). Characterization of melanin concentrating hormone in teleost hypothalamus. *Gen Comp Endocrinol* **65**, 423-431.

Kayal, M., Vercelloni, J., Lison de Loma, T., Bosserelle, P., Chancerelle, Y., Geoffroy, S., Stievenart, C., Michonneau, F., Penin, L., Planes, S. et al. (2012). Predator crown-of-thorns starfish (*Acanthaster planci*) outbreak, mass mortality of corals, and cascading effects on reef fish and benthic communities. *PLoS One* **7**, e47363.

Kim, D. K., Cho, E. B., Moon, M. J., Park, S., Hwang, J. I., Kah, O., Sower, S. A., Vaudry, H. and Seong, J. Y. (2011). Revisiting the evolution of gonadotropin-releasing hormones and their receptors in vertebrates: secrets hidden in genomes. *Gen Comp Endocrinol* **170**, 68-78.

Kim, Y. J., Zitnan, D., Galizia, C. G., Cho, K. H. and Adams, M. E. (2006). A command chemical triggers an innate behavior by sequential activation of multiple peptidergic ensembles. *Curr Biol* **16**, 1395-1407.

Kimura, S., Okada, M., Sugita, Y., Kanazawa, I. and Munekata, E. (1983). Novel neuropeptides, neurokinin alpha and beta isolated from porcine spinal cord. *Proc Jpn Acad Ser B* **59**, 101-104.

Kimura, T., Tanizawa, O., Mori, K., Brownstein, M. J. and Okayama, H. (1992). Structure and expression of a human oxytocin receptor. *Nature* **356**, 526-9.

Konturek, S. J., Zabielski, R., Konturek, J. W. and Czarnecki, J. (2003). Neuroendocrinology of the pancreas; role of brain-gut axis in pancreatic secretion. *Eur J Pharmacol* **481**, 1-14.

Kovac, J. R., Chrones, T., Preiksaitis, H. G. and Sims, S. M. (2006). Tachykinin receptor expression and function in human esophageal smooth muscle. *J Pharmacol Exp Ther* **318**, 513-520.

Kramer, S. J., Toschi, A., Miller, C. A., Kataoka, H., Quistad, G. B., Li, J. P., Carney, R. L. and Schooley, D. A. (1991). Identification of an allatostatin from the tobacco hornworm *Manduca sexta*. *Proc Natl Acad Sci U S A* **88**, 9458-9462.

Kurtz, M. M., Wang, R., Clements, M. K., Cascieri, M. A., Austin, C. P., Cunningham, B. R., Chicchi, G. G. and Liu, Q. (2002). Identification, localization and receptor characterization of novel mammalian substance P-like peptides. *Gene* **296**, 205-212.

Lang, T., Wacker, I., Steyer, J., Kaether, C., Wunderlich, I., Soldati, T., Gerdes, H. H. and Almers, W. (1997). Ca^{2+}-triggered peptide secretion in single cells imaged with green fluorescent protein and evanescent-wave microscopy. *Neuron* **18**, 857-863.

Lapraz, F., Rottinger, E., Duboc, V., Range, R., Duloquin, L., Walton, K., Wu, S. Y., Bradham, C., Loza, M. A., Hibino, T. et al. (2006). RTK and TGF-beta signaling pathways genes in the sea urchin genome. *Dev Biol* **300**, 132-152.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R. et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

Larsson, L. I. and Rehfeld, J. F. (1977). Evidence for a common evolutionary origin of gastrin and cholecystokinin. *Nature* **269**, 335-338.

Lebl, M., Hruby, V. J., Castrucci, A. M., Visconti, M. A. and Hadley, M. E. (1988). Melanin concentrating hormone analogues: contraction of the cyclic structure. 1. Agonist activity. *J Med Chem* **31**, 949-954.

Lechan, R. M., Wu, P., Jackson, I. M., Wolf, H., Cooperman, S., Mandel, G. and Goodman, R. H. (1986) Thyrotropin-releasing hormone precursor: characterization in rat brain. *Science*, **231,** 159-161.

Lee, C. M., Su, M. T. and Lee, H. J. (2009). Pigment dispersing factor: an output regulator of the circadian clock in the German cockroach. *J Biol Rhythms* **24**, 35-43.

Lee, J. H., Miele, M. E., Hicks, D. J., Phillips, K. K., Trent, J. M., Weissman, B. E. and Welch, D. R. (1996). KiSS-1, a novel human malignant melanoma metastasis-suppressor gene. *J Natl Cancer Inst* **88**, 1731-1737.

Li, B., Beeman, R. W. and Park, Y. (2011). Functions of duplicated genes encoding CCAP receptors in the red flour beetle, *Tribolium castaneum*. *J Insect Physiol* **57**, 1190-1197.

Li, R., Li, Y., Kristiansen, K. and Wang, J. (2008a). SOAP: short oligonucleotide alignment program. *Bioinformatics* **24**, 713-714.

Li, B., Predel, R., Neupert, S., Hauser, F., Tanaka, Y., Cazzamali, G., Williamson, M., Arakane, Y., Verleyen, P., Schoofs, L. et al. (2008b). Genomics, transcriptomics, and peptidomics of neuropeptides and protein hormones in the red flour beetle *Tribolium castaneum*. *Genome Res* **18**, 113-122.

Li, K. W., Holling, T., de With, N. D. and Geraerts, W. P. (1993). Purification and characterization of a novel tetradecapeptide that modulates oesophagus motility in *Lymnaea stagnalis*. *Biochem Biophys Res Commun* **197**, 1056-1061.

Li, Y., Hernandez-Martinez, S., Fernandez, F., Mayoral, J. G., Topalis, P., Priestap, H., Perez, M., Navare, A. and Noriega, F. G. (2006). Biochemical, molecular, and functional characterization of PISCF-allatostatin, a regulator of juvenile hormone biosynthesis in the mosquito *Aedes aegypti*. *J Biol Chem* **281**, 34048-34055.

Liddle, R. A. (1994) Cholecystokinin. In: *Gut Peptides* (ed. **Walsh, J. H., and Dockray, G.J**.). pp. 175–216. New York: Raven Press.

Light, A. and Du Vigneaud, V. (1958). On the nature of oxytocin and vasopressin from human pituitary. *Proc Soc Exp Biol Med* **98**, 692-696.

Lindemans, M., Janssen, T., Beets, I., Temmerman, L., Meelkop, E. and Schoofs, L. (2011). Gonadotropin-releasing hormone and adipokinetic hormone signaling systems share a common evolutionary origin. *Front Endocrinol (Lausanne)* **2**, 16.

Lindemans, M., Liu, F., Janssen, T., Husson, S. J., Mertens, I., Gade, G. and Schoofs, L. (2009). Adipokinetic hormone signaling through the gonadotropin-releasing hormone receptor modulates egg-laying in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* **106**, 1642-1647.

Liu, F., Baggerman, G., D'Hertog, W., Verleyen, P., Schoofs, L. and Wets, G. (2006). *In silico* identification of new secretory peptide genes in *Drosophila melanogaster*. *Mol Cell Proteomics* **5**, 510-522.

Liu, Y., Lu, D., Zhang, Y., Li, S., Liu, X. and Lin, H. (2010). The evolution of somatostatin in vertebrates. *Gene* **463**, 21-28.

Lloyd, P. E. and Connolly, C. M. (1989). Sequence of pedal peptide: a novel neuropeptide from the central nervous system of *Aplysia*. *J Neurosci* **9**, 312-317.

Loi, P. K., Emmal, S. A., Park, Y. and Tublitz, N. J. (2001). Identification, sequence and expression of a crustacean cardioactive peptide (CCAP) gene in the moth *Manduca sexta*. *J Exp Biol* **204**, 2803-2816.

Longley, R. D. and Peterman, M. (2013). Neuronal control of pedal sole cilia in the pond snail *Lymnaea stagnalis appressa*. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **199**, 71-86.

Lorenz, M. W., Kellner, R. and Hoffmann, K. H. (1995). Identification of two allatostatins from the cricket, *Gryllus bimaculatus de Geer* (*Ensifera*, *Gryllidae*): additional members of a family of neuropeptides inhibiting juvenile hormone biosynthesis. *Regul Pept* **57**, 227-236.

Lovejoy, D. A. (2005). Reproduction. In: *Neuroendocrinology*. pp. 275-306. Chicester: John Wiley and Sons, Ltd.

Lovejoy, D. A. and Balment, R. J. (1999). Evolution and physiology of the corticotropin-releasing factor (CRF) family of neuropeptides in vertebrates. *Gen Comp Endocrinol* **115**, 1-22.

Lovejoy, D. A. and Barsyte-Lovejoy, D. (2010). Characterization of a corticotropin-releasing factor (CRF)/diuretic hormone-like peptide from tunicates: insight into the origins of the vertebrate CRF family. *Gen Comp Endocrinol* **165**, 330-336.

Lovejoy, D. A. and Jahan, S. (2006). Phylogeny of the corticotropin-releasing factor family of peptides in the metazoa. *Gen Comp Endocrinol* **146**, 1-8.

Luo, C. W., Dewey, E. M., Sudo, S., Ewer, J., Hsu, S. Y., Honegger, H. W. and Hsueh, A. J. (2005). Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G-protein-coupled receptor LGR2. *Proc Natl Acad Sci U S A* **102**, 2820-2825.

Maestro, J. L., Aguilar, R., Pascual, N., Valero, M. L., Piulachs, M. D., Andreu, D., Navarro, I. and Belles, X. (2001). Screening of antifeedant activity in brain extracts led to the identification of sulfakinin as a satiety promoter in the German cockroach. Are arthropod sulfakinins homologous to vertebrate gastrins-cholecystokinins? *Eur J Biochem* **268**, 5824-5830.

Magnesen, T. and Redmond, K. (2012). Potential predation rates by the sea stars *Asterias rubens* and *Marthasterias glacialis*, on juvenile scallops, *Pecten maximus*, ready for sea ranching. *Aquac Int* **20**, 189-199.

Mains, R. E. and Eipper, B. A. (1999). Neuropeptide functions and regulation. In: *Basic Neurochemistry: Molecular, Cellular and Medical Aspects, 6th edition* (ed. **Siegel, G. J., Agranoff, B. W. and Albers, R. W., Fisher, S. K. and Uhler, M. D.**). Philadelphia: Lippincott-Raven.

Martínez, A., Riveros-Moreno, V., Polak, J. M., Moncada, S. and Sesma, P. (1994). Nitric oxide (NO) synthase immunoreactivity in the starfish *Marthasterias glacialis*. *Cell Tissue Res* **275**, 599-603.

Matsuda, K., Azuma, M. and Kang, K. S. (2012). Orexin system in teleost fish. *Vitam Horm* **89**, 341-361.

Mayer, A. G. (1909). On the use of magnesium in stupefying marine animals. *Biol Bull* **17**, 341-342.

McCall, C. and Singer, T. (2012). The animal and human neuroendocrinology of social cognition, motivation and behavior. *Nat Neurosci* **15**, 681-688.

McCormick, S. D. and Bradshaw, D. (2006). Hormonal control of salt and water balance in vertebrates. *Gen Comp Endocrinol* **147**, 3-8.

Melarange, R. and Elphick, M. R. (2003). Comparative analysis of nitric oxide and SALMFamide neuropeptides as general muscle relaxants in starfish. *J Exp Biol* **206**, 893-899.

Melarange, R., Potton, D. J., Thorndyke, M. C. and Elphick, M. R. (1999). SALMFamide neuropeptides cause relaxation and eversion of the cardiac stomach in starfish. *Proc R Soc B* **266**, 1785-1785.

Menschaert, G., Vandekerckhove, T. T., Baggerman, G., Landuyt, B., Sweedler, J. V., Schoofs, L., Luyten, W. and Van Criekinge, W. (2010). A hybrid, de novo based, genome-wide database search approach applied to the sea urchin neuropeptidome. *J Proteome Res* **9**, 990-996.

Meyer-Lindenberg, A., Domes, G., Kirsch, P. and Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nat Rev Neurosci* **12**, 524-38.

Michell, R. H., Kirk, C. J. and Billah, M. M. (1979). Hormonal stimulation of phosphatidylinositol breakdown with particular reference to the hepatic effects of vasopressin. *Biochem Soc Trans* **7**, 861-865.

Mirabeau, O. and Joly, J. S. (2013). Molecular evolution of peptidergic signaling systems in bilaterians. *Proc Natl Acad Sci U S A* **110**, e2028-37.

Mirabeau, O., Perlas, E., Severini, C., Audero, E., Gascuel, O., Possenti, R., Birney, E., Rosenthal, N. and Gross, C. (2007). Identification of novel peptide hormones in the human proteome by hidden Markov model screening. *Genome Res* **17**, 320-327.

Mita, M. (2013a). Relaxin-like gonad-stimulating substance in an echinoderm, the starfish: a novel relaxin system in reproduction of invertebrates. *Gen Comp Endocrinol* **181**, 241-245.

Mita, M. (2013b). Release of relaxin-like gonad-stimulating substance from starfish radial nerves by lonomycin. *Zoolog Sci* **30**, 602-606.

Mita, M., Oka, H., Thorndyke, M. C., Shibata, Y., Yoshikuni, M. and Nagahama, Y. (2004). Inhibitory effect of a SALMFamide neuropeptide on secretion of gonad-stimulating substance from radial nerves in the starfish *Asterina pectinifera*. *Zoolog Sci* **21**, 299-303.

Mita, M., Yoshikuni, M., Ohno, K., Shibata, Y., Paul-Prasanth, B., Pitchayawasin, S., Isobe, M. and Nagahama, Y. (2009). A relaxin-like peptide purified from radial nerves induces oocyte maturation and ovulation in the starfish, *Asterina pectinifera*. *Proc Natl Acad Sci U S A* **106**, 9507-9512.

Mohr, S. E., Hu, Y., Kim, K., Housden, B. E. and Perrimon, N. (2014). Resources for functional genomics studies in *Drosophila melanogaster*. *Genetics* **197**, 1-18.

Moore, S. J. and Thorndyke, M. C. (1993). Immunocytochemical mapping of the novel echinoderm neuropeptide SALMFamide 1 (S1) in the starfish *Asterias rubens*. *Cell Tissue Res* **274**, 605-618.

Morgan, K. and Millar, R. P. (2004). Evolution of GnRH ligand precursors and GnRH receptors in protochordate and vertebrate species. *Gen Comp Endocrinol* **139**, 191-197.

Moroz, L. L., Edwards, J. R., Puthanveettil, S. V., Kohn, A. B., Ha, T., Heyland, A., Knudsen, B., Sahni, A., Yu, F., Liu, L. et al. (2006). Neuronal transcriptome of *Aplysia*: neuronal compartments and circuitry. *Cell* **127**, 1453-1467.

Motokawa, T. (2011). Mechanical mutability in connective tissue of starfish body wall. *Biol Bull* **221**, 280-289.

Muschol, M. and Salzberg, B. M. (2000). Dependence of transient and residual calcium dynamics on action-potential patterning during neuropeptide secretion. *J Neurosci* **20**, 6773-6780.

Mutt, V. and Jorpes, J. E. (1968). Structure of porcine cholecystokinin-pancreozymin. 1. Cleavage with thrombin and with trypsin. *Eur J Biochem* **6**, 156-162.

Nachman, R. J., Giard, W., Favrel, P., Suresh, T., Sreekumar, S. and Holman, G. M. (1997). Insect myosuppressins and sulfakinins stimulate release of the digestive enzyme α-amylase in two invertebrates: the scallop *Pecten maximus* and insect *Rhynchophorus ferrugineus*. *Ann N Y Acad Sci* **814**, 335-338.

Nachman, R. J., Holman, G. M., Cook, B. J., Haddon, W. F. and Ling, N. (1986a). Leucosulfakinin-II, a blocked sulfated insect neuropeptide with homology to cholecystokinin and gastrin. *Biochem Biophys Res Commun* **140**, 357-364.

Nachman, R. J., Holman, G. M., Haddon, W. F. and Ling, N. (1986b). Leucosulfakinin, a sulfated insect neuropeptide with homology to gastrin and cholecystokinin. *Science* **234**, 71-73.

Nahon, J. L. (1994). The melanin-concentrating hormone: from the peptide to the gene. *Crit Rev Neurobiol* **8**, 221-262.

Nakabayashi, K., Matsumi, H., Bhalla, A., Bae, J., Mosselman, S., Hsu, S. Y. and Hsueh, A. J. W. (2002). Thyrostimulin, a heterodimer of two new human glycoprotein hormone subunits, activates the thyroid-stimulating hormone receptor. *J Clin Invest* **109**, 1445-1452.

Nässel, D. R. and Wegener, C. (2011). A comparative review of short and long neuropeptide F signaling in invertebrates: Any similarities to vertebrate neuropeptide Y signaling? *Peptides* **32**, 1335-1355.

Nathoo, A. N., Moeller, R. A., Westlund, B. A. and Hart, A. C. (2001). Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proc Natl Acad Sci U S A* **98**, 14000-14005.

Nawa, H., Kotani, H. and Nakanishi, S. (1984). Tissue-specific generation of two preprotachykinin mRNAs from one gene by alternative RNA splicing. *Nature* **312**, 729-734.

Newman, S. J., Elphick, M. R. and Thorndyke, M. C. (1995a). Tissue distribution of the SALMFamide neuropeptides S1 and S2 in the starfish *Asterias rubens* using novel monoclonal and polyclonal antibodies. II. Digestive system. *Proc Biol Sci* **261**, 187-192.

Newman, S. J., Elphick, M. R. and Thorndyke, M. C. (1995b). Tissue distribution of the SALMFamide neuropeptides S1 and S2 in the starfish *Asterias rubens* using novel monoclonal and polyclonal antibodies. I. Nervous and locomotory systems. *Proc Biol Sci* **261**, 139-145.

Niall, H. D., Keutmann, H. T., Copp, D. H. and Potts, J. T., Jr. (1969). Amino acid sequence of salmon ultimobranchial calcitonin. *Proc Natl Acad Sci U S A* **64**, 771-778.

Nichols, D. (1972). The water-vascular system in living and fossil echinoderms. *Paleontology*. **15,** 519-538.

Nichols, D. (1966). Functional morphology of the water-vascular system. In: *Physiology of Echinodermata* (ed. **Boolootian, R. A.**). pp. 219-244. New York: Wiley Interscience.

Nichols, R., Manoogian, B., Walling, E. and Mispelon, M. (2009). Plasticity in the effects of sulfated and nonsulfated sulfakinin on heart contractions. *Front Biosci (Landmark Ed)* **14**, 4035-4043.

Nieto, J., Veelaert, D., Derua, R., Waelkens, E., Cerstiaens, A., Coast, G., Devreese, B., Van Beeumen, J., Calderon, J., De Loof, A. et al. (1998). Identification of one tachykinin- and two kinin-related peptides in the brain of the white shrimp, *Penaeus vannamei*. *Biochem Biophys Res Commun* **248**, 406-411.

Nordstrom, K. J., Fredriksson, R. and Schioth, H. B. (2008). The amphioxus (*Branchiostoma floridae*) genome contains a highly diversified set of G-protein-coupled receptors. *BMC Evol Biol* **8**, 9.

O'Hara, T. D., Hugall, A. F., Thuy, B. and Moussalli, A. (2014). Phylogenomic resolution of the class Ophiuroidea unlocks a global microfossil record. *Curr Biol* **24**, 1874-1879.

O'Neill, P. (1989). Structure and mechanics of starfish body wall. *J Exp Biol* **147**, 53-89.

O'Shea, M. and Rayne, R. C. (1992). Adipokinetic hormones: cell and molecular biology. *Experientia* **48**, 430-438.

Oakley, A. E., Clifton, D. K. and Steiner, R. A. (2009). Kisspeptin signaling in the brain. *Endocr Rev* **30**, 713-743.

Ocampo Daza, D., Sundstrom, G., Bergqvist, C. A. and Larhammar, D. (2012). The evolution of vertebrate somatostatin receptors and their gene regions involves extensive chromosomal rearrangements. *BMC Evol Biol* **12**, 231.

Ochiai, H., Fujita, K., Suzuki, K., Nishikawa, M., Shibata, T., Sakamoto, N. and Yamamoto, T. (2010). Targeted mutagenesis in the sea urchin embryo using zinc-finger nucleases. *Genes Cells* **15**, 875-885.

Ogoshi, M., Inoue, K., Naruse, K. and Takei, Y. (2006). Evolutionary history of the calcitonin gene-related peptide family in vertebrates revealed by comparative genomic analyses. *Peptides* **27**, 3154-3164.

Ohtani, M., Iwakoshi, E., Muneoka, Y., Minakata, H. and Nomoto, K. (1999). Isolation and characterisation of bioactive peptides from the sea cucumber, *Stichopus japonicus*. In: *Peptide Science – Present and Future* (ed. **Shimonishi, Y.**). pp. 419-420. Dordrecht: Kluwer Academic Publishers.

Okamura, N., Hashimoto, K., Iyo, M., Shimizu, E., Dempfle, A., Friedel, S. and Reinscheid, R. K. (2007). Gender-specific association of a functional coding polymorphism in the neuropeptide S receptor gene with panic disorder but not with schizophrenia or attention-deficit/hyperactivity disorder. *Prog Neuropsychopharmacol Biol Psychiatry* **31**, 1444-1448.

Oldham, W. M. and Hamm, H. E. (2008). Heterotrimeric G-protein activation by G-protein-coupled receptors. *Nat Rev Mol Cell Biol* **9**, 60-71.

Otara, C. B., Jones, C. E., Younan, N. D., Viles, J. H. and Elphick, M. R. (2014). Structural analysis of the starfish SALMFamide neuropeptides S1 and S2: the N-terminal region of S2 facilitates self-association. *Biochim Biophys Acta* **1844**, 358-365.

Oumi, T., Ukena, K., Matsushima, O., Ikeda, T., Fujita, T., Minakata, H. and Nomoto, K. (1996). Annetocin, an annelid oxytocin-related peptide, induces egg-laying behavior in the earthworm, *Eisenia foetida*. *J Exp Zool* **276**, 151-156.

Page, N. M., Bell, N. J., Gardiner, S. M., Manyonda, I. T., Brayley, K. J., Strange, P. G. and Lowry, P. J. (2003). Characterization of the endokinins: human tachykinins with cardiovascular activity. *Proc Natl Acad Sci U S A* **100**, 6245-6250.

Park, J. H., Schroeder, A. J., Helfrich-Forster, C., Jackson, F. R. and Ewer, J. (2003). Targeted ablation of CCAP neuropeptide-containing neurons of *Drosophila* causes specific defects in execution and circadian timing of ecdysis behavior. *Development* **130**, 2645-2656.

Park, J. I., Semyonov, J., Chang, C. L. and Hsu, S. Y. (2005). Conservation of the heterodimeric glycoprotein hormone subunit family proteins and the LGR signaling system from nematodes to humans. *Endocrine* **26**, 267-276.

Park, Y., Kim, Y. J. and Adams, M. E. (2002). Identification of G protein-coupled receptors for *Drosophila* PRXamide peptides, CCAP, corazonin, and AKH supports a theory of ligand-receptor coevolution. *Proc Natl Acad Sci U S A* **99**, 11423-11428.

Patruno, M., Thorndyke, M. C., Candia Carnevali, M. D., Bonasoro, F. and Beesley, P. W. (2001). Growth factors, heat-shock proteins and regeneration in echinoderms. *J Exp Biol* **204**, 843-848.

Perillo, M. and Arnone, M. I. (2014). Characterization of insulin-like peptides (ILPs) in the sea urchin *Strongylocentrotus purpuratus*: insights on the evolution of the insulin family. *Gen Comp Endocrinol* **205**, 68-79.

Philippe, H., Brinkmann, H., Copley, R. R., Moroz, L. L., Nakano, H., Poustka, A. J., Wallberg, A., Peterson, K. J. and Telford, M. J. (2011). Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. *Nature* **470**, 255-258.

Phlippen, M. K., Webster, S. G., Chung, J. S. and Dircksen, H. (2000). Ecdysis of decapod crustaceans is associated with a dramatic release of crustacean cardioactive peptide into the haemolymph. *J Exp Biol* **203**, 521-536.

Pielecka-Fortuna, J., Chu, Z. and Moenter, S. M. (2008). Kisspeptin acts directly and indirectly to increase gonadotropin-releasing hormone neuron activity and its effects are modulated by estradiol. *Endocrinology* **149**, 1979-1986.

Pierce, J. G. and Parsons, T. F. (1981). Glycoprotein hormones: structure and function. *Annu Rev Biochem* **50**, 465-495.

Pitti, T. and Manoj, N. (2012). Molecular evolution of the neuropeptide S receptor. *PLoS One* **7**, e34046.

Pradayrol, L., Chayvialle, J. A., Carlquist, M. and Mutt, V. (1978). Isolation of a porcine intestinal peptide with C-terminal somatostatin. *Biochem Biophys Res Commun* **85**, 701-708.

Pradayrol, L., Jornvall, H., Mutt, V. and Ribet, A. (1980). N-terminally extended somatostatin: the primary structure of somatostatin-28. *FEBS Lett* **109**, 55-58.

Pratt, G. E., Farnsworth, D. E., Fok, K. F., Siegel, N. R., McCormack, A. L., Shabanowitz, J., Hunt, D. F. and Feyereisen, R. (1991). Identity of a second type of allatostatin from cockroach brains: an octadecapeptide amide with a tyrosine-rich address sequence. *Proc Natl Acad Sci U S A* **88**, 2412-2416.

Pratt, G. E., Farnsworth, D. E., Siegel, N. R., Fok, K. F. and Feyereisen, R. (1989). Identification of an allatostatin from adult *Diploptera punctata*. *Biochem Biophys Res Commun* **163**, 1243-1247.

Price, M. D., Merte, J., Nichols, R., Koladich, P. M., Tobe, S. S. and Bendena, W. G. (2002). *Drosophila melanogaster* flatline encodes a myotropin orthologue to *Manduca sexta* allatostatin. *Peptides* **23**, 787-794.

Priyam, A., Woodcroft, B. J, Rai, V. and Wurm, Y. (*in prep*). SequenceServer: BLAST searching made easy. Available at: http://sequenceserver.com.

Proux, J. P., Miller, C. A., Li, J. P., Carney, R. L., Girardie, A., Delaage, M. and Schooley, D. A. (1987). Identification of an arginine vasopressin-like diuretic hormone from *Locusta migratoria*. *Biochem Biophys Res Commun* **149**, 180-186.

Putnam, N. H., Butts, T., Ferrier, D. E., Furlong, R. F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J. K. et al. (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**, 1064-1071.

Rao, K. R. and Riehm, J. P. (1993). Pigment-dispersing hormones. *Ann N Y Acad Sci* **680**, 78-88.

Rapoport, T. A., Jungnickel, B. and Kutay, U. (1996). Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu Rev Biochem* **65**, 271-303.

Reich, G. (1992). A new peptide of the oxytocin/vasopressin family isolated from nerves of the cephalopod *Octopus vulgaris*. *Neurosci Lett* **134**, 191-194.

Reinecke, M. and Collet, C. (1998). The phylogeny of the insulin-like growth factors. *Int Rev Cytol* **183**, 1-94.

Reinscheid, R. K. (2007). Phylogenetic appearance of neuropeptide S precursor proteins in tetrapods. *Peptides* **28**, 830-837.

Rholam, M., Brakch, N., Germain, D., Thomas, D. Y., Fahy, C., Boussetta, H., Boileau, G. and Cohen, P. (1995). Role of amino acid sequences flanking dibasic cleavage sites in precursor proteolytic processing. The importance of the first residue C-terminal of the cleavage site. *Eur J Biochem* **227**, 707-714.

Richard, N., Corvaisier, S., Camacho, E. and Kottler, M. L. (2009). KiSS-1 and GPR54 at the pituitary level: overview and recent insights. *Peptides* **30**, 123-129.

Richards, S. Gibbs, R. A. Weinstock, G. M. Brown, S. J. Denell, R. Beeman, R. W. Gibbs, R. Beeman, R. W. Brown, S. J. Bucher, G. et al. (2008). The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**, 949-955.

Richter, K., Kawashima, E., Egger, R. and Kreil, G. (1984). Biosynthesis of thyrotropin releasing hormone in the skin of *Xenopus laevis*: partial sequence of the precursor deduced from cloned cDNA. *EMBO J* **3**, 617–621.

Rivier, J., Brazeau, P., Vale, W. and Guillemin, R. (1975). Somatostatin analogs. Relative importance of the disulfide bridge and of the Ala-Gly side chain for biological activity. *J Med Chem* **18**, 123-126.

Roberts, M. P. and Campbell, A. C. (1988) Functional anatomy of pedicellariae from *Asterias rubens* L. In: *Echinoderm biology* (ed. **Burke, R. D., Mladenov, P.V., Lambert, P. and Parsley, R. L**.). pp. 725-733. Rotterdam: Balkema.

Roch, G. J., Busby, E. R. and Sherwood, N. M. (2011). Evolution of GnRH: diving deeper. *Gen Comp Endocrinol* **171**, 1-16.

Roch, G. J., Tello, J. A. and Sherwood, N. M. (2014). At the transition from invertebrates to vertebrates, a novel GnRH-like peptide emerges in amphioxus. *Mol Biol Evol* **31**, 765-778.

Roth, A. L., Marzola, E., Rizzi, A., Arduin, M., Trapella, C., Corti, C., Vergura, R., Martinelli, P., Salvadori, S., Regoli, D. et al. (2006). Structure-activity studies on neuropeptide S: identification of the amino acid residues crucial for receptor activation. *J Biol Chem* **281**, 20809-20816.

Rouille, Y., Duguay, S. J., Lund, K., Furuta, M., Gong, Q., Lipkind, G., Oliva, A. A., Jr., Chan, S. J. and Steiner, D. F. (1995). Proteolytic processing mechanisms in the biosynthesis of neuroendocrine peptides: the subtilisin-like proprotein convertases. *Front Neuroendocrinol* **16**, 322-361.

Rowe, M. L., Achhala, S. and Elphick, M. R. (2014). Neuropeptides and polypeptide hormones in echinoderms: new insights from analysis of the transcriptome of the sea cucumber *Apostichopus japonicus*. *Gen Comp Endocrinol* **197**, 43-55.

Rowe, M. L. and Elphick, M. R. (2010). Discovery of a second SALMFamide gene in the sea urchin *Strongylocentrotus purpuratus* reveals that L-type and F-type SALMFamide neuropeptides coexist in an echinoderm species. *Mar Genomics* **3**, 91-97.

Rowe, M. L. and Elphick, M. R. (2012). The neuropeptide transcriptome of a model echinoderm, the sea urchin *Strongylocentrotus purpuratus*. *Gen Comp Endocrinol* **179**, 331-344.

Ruppert, E. E., Fox, R. S. and Barnes, R.D. (2004). Invertebrate Zoology, 7th edition. Belmont, CA: Thomson, Brooks/Cole.

Saha, A. K., Tamori, M., Inoue, M., Nakajima, Y. and Motokawa, T. (2006). NGIWYamide-induced contraction of tube feet and distribution of NGIWYamide-like immunoreactivity in nerves of the starfish *Asterina pectinifera*. *Zoolog Sci* **23**, 627-632.

Sakurai, T. (2007). The neural circuit of orexin (hypocretin): maintaining sleep and wakefulness. *Nat Rev Neurosci* **8**, 171-181.

Sakurai, T., Amemiya, A., Ishii, M., Matsuzaki, I., Chemelli, R. M., Tanaka, H., Williams, S. C., Richarson, J. A., Kozlowski, G. P., Wilson, S. et al. (1998). Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* **92**, 573-585.

Salzet, M., Verger-Bocquet, M., Vandenbulcke, F. and Van Minnen, J. (1997). Leech egg-laying-like hormone: structure, neuronal distribution and phylogeny. *Brain Res Mol Brain Res* **49**, 211-221.

Samanta, M. P., Tongprasit, W., Istrail, S., Cameron, R. A., Tu, Q., Davidson, E. H. and Stolc, V. (2006). The transcriptome of the sea urchin embryo. *Science* **314**, 960-962.

Satake, H., Ogasawara, M., Kawada, T., Masuda, K., Aoyama, M., Minakata, H., Chiba, T., Metoki, H., Satou, Y. and Satoh, N. (2004). Tachykinin and tachykinin receptor of an ascidian, *Ciona intestinalis*: evolutionary origin of the vertebrate tachykinin family. *J Biol Chem* **279**, 53798-53805.

Sawyer, W. H. (1977). Evolution of neurohypophyseal hormones and their receptors. *Fed Proc* **36**, 1842-1847.

Scheller, R. H., Jackson, J. F., McAllister, L. B., Rothman, B. S., Mayeri, E. and Axel, R. (1983). A single gene encodes multiple neuropeptides mediating a stereotyped behavior. *Cell* **32**, 7-22.

Schoofs, L. and Nachman, R. J. (2006) Sulfakinins. In: *Handbook of biologically active peptides* (ed. **Kastin, A.J.**). pp. 183-187. Amsterdam: Elsevier.

Schoofs, L., Janssen, T. and Nachman, R. J. (2013). Sulfakinins. In: *Handbook of biologically active peptides, vol. 2* (ed. **Kastin, A. J.**). pp. 310–314. Amsterdam: Elsevier.

Seidah, N. G. and Chretien, M. (1999). Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. *Brain Res* **848**, 45-62.

Sekiguchi, T., Suzuki, N., Fujiwara, N., Aoyama, M., Kawada, T., Sugase, K., Murata, Y., Sasayama, Y., Ogasawara, M. and Satake, H. (2009). Calcitonin in a protochordate, *Ciona intestinalis* - the prototype of the vertebrate calcitonin/calcitonin gene-related peptide superfamily. *FEBS Journal* **276**, 4437-4447.

Sellami, A., Agricola, H. J. and Veenstra, J. A. (2011). Neuroendocrine cells in *Drosophila melanogaster* producing GPA2/GPB5, a hormone with homology to LH, FSH and TSH. *Gen Comp Endocrinol* **170**, 582-588.

Seminara, S. B., Messager, S., Chatzidaki, E. E., Thresher, R. R., Acierno, J. S., Jr., Shagoury, J. K., Bo-Abbas, Y., Kuohung, W., Schwinof, K. M., Hendrick, A. G. et al. (2003). The GPR54 gene as a regulator of puberty. *N Engl J Med* **349**, 1614-1627.

Severini, C., Improta, G., Falconieri-Erspamer, G., Salvadori, S. and Erspamer, V. (2002). The tachykinin peptide family. *Pharmacol Rev* **54**, 285-322.

Shibusawa, N., Hashimoto, K. and Yamada, M. (2008). Thyrotropin-releasing hormone (TRH) in the cerebellum. *Cerebellum* **7**, 84-95.

Shimizu, Y., Matsuyama, H., Shiina, T., Takewaki, T. and Furness, J. B. (2008). Tachykinins and their functions in the gastrointestinal tract. *Cell Mol Life Sci* **65**, 295-311.

Simakov, O., Marletaz, F., Cho, S. J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D. H., Larsson, T., Lv, J., Arendt, D. et al. (2013). Insights into bilaterian evolution from three spiralian genomes. *Nature* **493**, 526-531.

Simon, M. I., Strathmann, M. P. and Gautam, N. (1991). Diversity of G-proteins in signal transduction. *Science* **252**, 802-808.

Siviter, R. J., Coast, G. M., Winther, A. M., Nachman, R. J., Taylor, C. A., Shirras, A. D., Coates, D., Isaac, R. E. and Nässel, D. R. (2000). Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian preprotachykinin A. *J Biol Chem* **275**, 23273-23280.

Smith, J. E. (1965). Echinodermata. In: *Structure and Function in the Nervous Systems of Invertebrates* (ed. **Bullock, T. H. and Horridge, G. A.**). pp. 1519–1558. London: W. H. Freeman and Co.

Sodergren, E., Weinstock, G. M., Davidson, E. H., Cameron, R. A., Gibbs, R. A., Angerer, R. C., Angerer, L. M., Arnone, M. I., Burgess, D. R., Burke, R. D. et al. (2006). The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* **314**, 941-952.

Soehler, S., Stengl, M. and Reischig, T. (2011). Circadian pacemaker coupling by multi-peptidergic neurons in the cockroach *Leucophaea maderae*. *Cell Tissue Res* **343**, 559-577.

Sower, S. A., Freamat, M. and Kavanaugh, S. I. (2009). The origins of the vertebrate hypothalamic-pituitary-gonadal (HPG) and hypothalamic-pituitary-thyroid (HPT) endocrine systems: new insights from lampreys. *Gen Comp Endocrinol* **161**, 20-29.

Stafflinger, E., Hansen, K. K., Hauser, F., Schneider, M., Cazzamali, G., Williamson, M. and Grimmelikhuijzen, C. J. (2008). Cloning and identification of an oxytocin/vasopressin-like receptor and its ligand from insects. *Proc Natl Acad Sci U S A* **105**, 3262-3267.

Stangier, J., Hilbich, C., Beyreuther, K. and Keller, R. (1987). Unusual cardioactive peptide (CCAP) from pericardial organs of the shore crab *Carcinus maenas*. *Proc Natl Acad Sci U S A* **84**, 575-579.

Stangier, J., Hilbich, C., Burdzik, S. and Keller, R. (1992). Orcokinin: a novel myotropic peptide from the nervous system of the crayfish, *Orconectes limosus*. *Peptides* **13**, 859-864.

Steiner, D. F. (1998). The proprotein convertases. *Curr Opin Chem Biol* **2**, 31-39.

Stoop, R. (2012). Neuromodulation by oxytocin and vasopressin. *Neuron* **76**, 142-159.

Strand, F. L. (1999). Neuropeptides: regulators of physiological processes. Cambridge, MA: MIT Press.

Studer, R. O., Trzeciak, A. and Lergier, W. (1973). Isolierung und Aminosäuresequenz von Substanz P aus Pferdedarm. *Helvetica Chimica Acta* **56**, 860-866.

Sudhof, T. C. (2012). Calcium control of neurotransmitter release. *Cold Spring Harb Perspect Biol* **4**, a011353.

Sudo, S., Kuwabara, Y., Park, J. I., Hsu, S. Y. and Hsueh, A. J. (2005). Heterodimeric fly glycoprotein hormone-alpha2 (GPA2) and glycoprotein hormone-beta5 (GPB5) activate fly leucine-rich repeat-containing G protein-coupled receptor-1 (DLGR1) and stimulation of human thyrotropin receptors by chimeric fly GPA2 and human GPB5. *Endocrinology* **146**, 3596-3604.

Sugimoto, T., Saito, M., Mochizuki, S., Watanabe, Y., Hashimoto, S. and Kawashima, H. (1994). Molecular cloning and functional expression of a cDNA encoding the human V1b vasopressin receptor. *J Biol Chem* **269**, 27088-27092.

Taghert, P. H. and Nitabach, M. N. (2012). Peptide neuromodulation in invertebrate model systems. *Neuron* **76**, 82-97.

Telford, M. J. (2006). Animal phylogeny. *Curr Biol* **16**, R981-5.

Telford, M. J., Lowe, C. J., Cameron, C. B., Ortega-Martinez, O., Aronowicz, J., Oliveri, P. and Copley, R. R. (2014). Phylogenomic analysis of echinoderm class relationships supports Asterozoa. *Proc Biol Sci* **281**, 20140479.

Tensen, C. P., Cox, K. J., Smit, A. B., van der Schors, R. C., Meyerhof, W., Richter, D., Planta, R. J., Hermann, P. M., van Minnen, J., Geraerts, W. P. et al. (1998). The *lymnaea* cardioexcitatory peptide (LyCEP) receptor: a G-protein-coupled receptor for a novel member of the RFamide neuropeptide family. *J Neurosci* **18**, 9812-9821.

Thibonnier, M., Auzan, C., Madhun, Z., Wilkins, P., Berti-Mattera, L. and Clauser, E. (1994). Molecular cloning, sequencing, and functional expression of a cDNA encoding the human V1a vasopressin receptor. *J Biol Chem* **269**, 3304-3310.

Thorndyke, M. C., Chen, W. C., Beesley, P. W. and Patruno, M. (2001). Molecular approach to echinoderm regeneration. *Microsc Res Tech* **55**, 474-485.

Thorsell, A. and Heilig, M. (2002). Diverse functions of neuropeptide Y revealed using genetically modified animals. *Neuropeptides* **36**, 182-193.

Timmers, M. A., Bird, C. E., Skillings, D. J., Smouse, P. E. and Toonen, R. J. (2012). There's no place like home: crown-of-thorns outbreaks in the central pacific are regionally derived and independent events. *PLoS One* **7**, e31159.

Tooze, S. A. and Huttner, W. B. (1990). Cell-free protein sorting to the regulated and constitutive secretory pathways. *Cell* **60**, 837-847.

Torres-Lugo, M. and Peppas, N. A. (2000). Transmucosal delivery systems for calcitonin: a review. *Biomaterials* **21**, 1191-1196.

Tostivint, H., Lihrmann, I. and Vaudry, H. (2008). New insight into the molecular evolution of the somatostatin family. *Mol Cell Endocrinol* **286**, 5-17.

Tsai, P. S. (2006). Gonadotropin-releasing hormone in invertebrates: structure, function, and evolution. *Gen Comp Endocrinol* **148**, 48-53.

Tsai, P. S. and Zhang, L. (2008). The emergence and loss of gonadotropin-releasing hormone in protostomes: orthology, phylogeny, structure, and function. *Biol Reprod* **79**, 798-805.

Tsujino, N. and Sakurai, T. (2009). Orexin/hypocretin: a neuropeptide at the interface of sleep, energy homeostasis, and reward system. *Pharmacol Rev* **61**, 162-176.

Tu, Q., Cameron, R. A., Worley, K. C., Gibbs, R. A. and Davidson, E. H. (2012). Gene structure in the sea urchin *Strongylocentrotus purpuratus* based on transcriptome analysis. *Genome Res* **22**, 2079-2087.

Ukena, K., Iwakoshi-Ukena, E. and Hikosaka, A. (2008). Unique form and osmoregulatory function of a neurohypophysial hormone in a urochordate. *Endocrinology* **149**, 5254-5261.

Ukena, K., Oumi, T., Matsushima, O., Ikeda, T., Fujita, T., Minakata, H. and Nomoto, K. (1995). A novel gut tetradecapeptide isolated from the earthworm, *Eisenia foetida*. *Peptides* **16**, 995-999.

Urano, A., Hyodo, S. and Suzuki, M. (1992). Molecular evolution of neurohypophysial hormone precursors. *Prog Brain Res* **92**, 39-46.

von Euler, U. S. and Gaddum, J. H. (1931). An unidentified depressor substance in certain tissue extracts. *J Physiol* **72**, 74-87.

Valsalan, R. and Manoj, N. (2014). Evolutionary history of the neuropeptide S receptor/neuropeptide S system. *Gen Comp Endocrinol* **209**, 11-20.

van den Pol, A. N. (2012). Neuropeptide transmission in brain circuits. *Neuron* **76**, 98-115.

van Kesteren, R. E. and Geraerts, W. P. (1998). Molecular evolution of ligand-binding specificity in the vasopressin/oxytocin receptor family. *Ann N Y Acad Sci* **839**, 25-34.

van Kesteren, R. E., Smit, A. B., De Lange, R. P., Kits, K. S., Van Golen, F. A., Van Der Schors, R. C., De With, N. D., Burke, J. F. and Geraerts, W. P. (1995). Structural and functional evolution of the vasopressin/oxytocin superfamily: vasopressin-related conopressin is the only member present in *Lymnaea*, and is involved in the control of sexual behavior. *J Neurosci* **15**, 5989-5998.

van Kesteren, R. E., Smit, A. B., de With, N. D., van Minnen, J., Dirks, R. W., van der Schors, R. C. and Joosse, J. (1992). A vasopressin-related peptide in the mollusc *Lymnaea stagnalis*: peptide structure, prohormone organization, evolutionary and functional aspects of *Lymnaea* conopressin. *Prog Brain Res* **92**, 47-57.

van Kesteren, R. E., Tensen, C. P., Smit, A. B., van Minnen, J., Kolakowski, L. F., Meyerhof, W., Richter, D., van Heerikhuizen, H., Vreugdenhil, E. and Geraerts, W. P. (1996). Co-evolution of ligand-receptor pairs in the vasopressin/oxytocin superfamily of bioactive peptides. *J Biol Chem* **271**, 3619-3626.

Van Loy, T., Van Hiel, M. B., Vandersmissen, H. P., Poels, J., Mendive, F., Vassart, G. and Vanden Broeck, J. (2007). Evolutionary conservation of bursicon in the animal kingdom. *Gen Comp Endocrinol* **153**, 59-63.

Van Loy, T., Vandersmissen, H. P., Poels, J., Van Hiel, M. B., Verlinden, H. and Vanden Broeck, J. (2010). Tachykinin-related peptides and their receptors in invertebrates: a current view. *Peptides* **31**, 520-524.

Vanden Broeck, J. (2001). Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster*. *Peptides* **22**, 241-254.

Varro, A. and Dockray, G. J. (1993). Post-translational processing of progastrin: inhibition of cleavage, phosphorylation and sulphation by brefeldin A. *Biochem J* **295**, 813-819.

Veelaert, D., Baggerman, G., Derua, R., Waelkens, E., Meeusen, T., Vande Water, G., De Loof, A. and Schoofs, L. (1999). Identification of a new tachykinin from the midgut of the desert locust, *Schistocerca gregaria*, by ESI-Qq-oa-TOF mass spectrometry. *Biochem Biophys Res Commun* **266**, 237-242.

Veenstra, J. A. (2000). Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. *Arch Insect Biochem Physiol* **43**, 49-63.

Veenstra, J. A. (2009). Allatostatin C and its paralog allatostatin double C: the arthropod somatostatins. *Insect Biochem Mol Biol* **39**, 161-170.

Veenstra, J. A. (2010). Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *Gen Comp Endocrinol* **167**, 86-103.

Veenstra, J. A. (2011). Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. *Gen Comp Endocrinol* **171**, 160-175.

Veenstra, J. A., Lehman, H. K. and Davis, N. T. (1994). Allatotropin is a cardioacceleratory peptide in *Manduca sexta*. *J Exp Biol* **188**, 347-354.

Viollet, C., Lepousez, G., Loudes, C., Videau, C., Simon, A. and Epelbaum, J. (2008). Somatostatinergic systems in brain: networks and functions. *Mol Cell Endocrinol* **286**, 75-87.

von Hehn, G. (1970). The fine structure of the hyponeural nerve system on the starfish (*Asterias rubens* L.). *Z Zellforsch Mikrosk Anat* **105**, 137-154.

von Heijne, G. (1985). Signal sequences. The limits of variation. *J Mol Biol* **184**, 99-105.

Vreugdenhil, E., Jackson, J. F., Bouwmeester, T., Smit, A. B., Van Minnen, J., Van Heerikhuizen, H., Klootwijk, J. and Joosse, J. (1988). Isolation, characterization, and evolutionary aspects of a cDNA clone encoding multiple neuropeptides involved in the stereotyped egg-laying behavior of the freshwater snail *Lymnaea stagnalis*. *J Neurosci* **8**, 4184-4191.

Vuerinckx, K., Verlinden, H., Lindemans, M., Broeck, J. V. and Huybrechts, R. (2011). Characterization of an allatotropin-like peptide receptor in the red flour beetle, *Tribolium castaneum*. *Insect Biochem Mol Biol* **41**, 815-822.

Walsh, D. A. and McWilliams, D. F. (2006). Tachykinins and the cardiovascular system. *Curr Drug Targets* **7**, 1031-1042.

Weaver, R. J. and Audsley, N. (2009). Neuropeptide regulators of juvenile hormone synthesis: structures, functions, distribution, and unanswered questions. *Ann N Y Acad Sci* **1163**, 316-329.

Webster, S. G. (1991). Amino acid sequence of putative moult-inhibiting hormone from the crab *Carcinus maenas*. *Proc Biol Sci* **244**, 247-252.

Webster, S. G., Wilcockson, D. C., Mrinalini and Sharp, J. H. (2013). Bursicon and neuropeptide cascades during the ecdysis program of the shore crab, *Carcinus maenas*. *Gen Comp Endocrinol* **182**, 54-64.

Wei, H. and Stengl, M. (2011). Light affects the branching pattern of peptidergic circadian pacemaker neurons in the brain of the cockroach *Leucophaea maderae*. *J Biol Rhythms* **26**, 507-517.

Wei, Z., Baggerman, G., J. Nachman, R., Goldsworthy, G., Verhaert, P., De Loof, A. and Schoofs, L. (2000). Sulfakinins reduce food intake in the desert locust, *Schistocerca gregaria*. *J Insect Physiol* **46**, 1259-1265.

Wendelaar Bonga, S. E. and Pang, P. K. (1991). Control of calcium regulating hormones in the vertebrates: parathyroid hormone, calcitonin, prolactin, and stanniocalcin. *Int Rev Cytol* **128**, 139-213.

Wilcockson, D. C. and Webster, S. G. (2008). Identification and developmental expression of mRNAs encoding putative insect cuticle hardening hormone, bursicon in the green shore crab *Carcinus maenas*. *Gen Comp Endocrinol* **156**, 113-125.

Wilkie, I. C., Griffiths, G. V. R. and Glennie, S. F. (1990). Morphological and physiological aspects of the autotomy plane in the aboral integument of *Asterias rubens* L. (Echinodermata). In: *Echinoderm Research* (ed. **De Ridder, C., Dubois, P., Marie-Christine, L. and Jangoux, M.**). pp. 301-313. Rotterdam: Balkema.

Wilkie, I. C. (2001). Autotomy as a prelude to regeneration in echinoderms. *Microsc Res Tech* **55**, 369-396.

Wilkie, I. C. (2005). Mutable collagenous tissue: overview and biotechnological perspective. *Prog Mol Subcell Biol* **39**, 221-250.

Wilkinson, T. N., Speed, T. P., Tregear, G. W. and Bathgate, R. A. (2005). Evolution of the relaxin-like peptide family. *BMC Evol Biol*, **5,** 14.

Wilkinson, T. N. and Bathgate, R. A. (2007). The evolution of the relaxin peptide family and their receptors. *Adv Exp Med Biol* **612**, 1-13.

Williamson, M., Lenz, C., Winther, A. M., Nässel, D. R. and Grimmelikhuijzen, C. J. (2001). Molecular cloning, genomic organization, and expression of a C-type (*Manduca sexta*-type) allatostatin preprohormone from *Drosophila melanogaster*. *Biochem Biophys Res Commun* **282**, 124-130.

Wilson, S. and Falkmer, S. (1965). Starfish insulin. *Can J Biochem* **43**, 1615-1624.

Winslow, J. T., Hastings, N., Carter, C. S., Harbaugh, C. R. and Insel, T. R. (1993). A role for central vasopressin in pair bonding in monogamous prairie voles. *Nature* **365**, 545-548.

Wistrand, M., Kall, L. and Sonnhammer, E. L. (2006). A general model of G protein-coupled receptor sequences and its application to detect remote homologs. *Protein Sci* **15**, 509-521.

Woodhead, A. P., Stay, B., Seidel, S. L., Khan, M. A. and Tobe, S. S. (1989). Primary structure of four allatostatins: neuropeptide inhibitors of juvenile hormone synthesis. *Proc Natl Acad Sci U S A* **86**, 5997-6001.

Xu, Y. L., Reinscheid, R. K., Huitron-Resendiz, S., Clark, S. D., Wang, Z., Lin, S. H., Brucher, F. A., Zeng, J., Ly, N. K., Henriksen, S. J. et al. (2004). Neuropeptide S: a neuropeptide promoting arousal and anxiolytic-like effects. *Neuron* **43**, 487-497.

Yamanaka, N., Roller, L., Zitnan, D., Satake, H., Mizoguchi, A., Kataoka, H. and Tanaka, Y. (2011). *Bombyx* orcokinins are brain-gut peptides involved in the neuronal regulation of ecdysteroidogenesis. *J Comp Neurol* **519**, 238-246.

Yamanaka, N., Yamamoto, S., Zitnan, D., Watanabe, K., Kawada, T., Satake, H., Kaneko, Y., Hiruma, K., Tanaka, Y., Shinoda, T. et al. (2008). Neuropeptide receptor transcriptome reveals unidentified neuroendocrine pathways. *PLoS One* **3**, e3048.

Yegorov, S. and Good, S. (2012). Using paleogenomics to study the evolution of gene families: origin and duplication history of the relaxin family hormones and their receptors. *PLoS One* **7**, e32923.

Young, L. J. and Wang, Z. (2004). The neurobiology of pair bonding. *Nat Neurosci* **7**, 1048-1054.

Zhang, C., Roepke, T. A., Kelly, M. J. and Ronnekleiv, O. K. (2008). Kisspeptin depolarizes gonadotropin-releasing hormone neurons through activation of TRPC-like cationic channels. *J Neurosci* **28**, 4423-4434.

Zhang, L., Wayne, N. L., Sherwood, N. M., Postigo, H. R. and Tsai, P. S. (2000a). Biological and immunological characterization of multiple GnRH in an opisthobranch mollusk, *Aplysia californica*. *Gen Comp Endocrinol* **118**, 77-89.

Zhang, Y., Lu, L., Furlonger, C., Wu, G. E. and Paige, C. J. (2000b). Hemokinin is a hematopoietic-specific tachykinin that regulates B lymphopoiesis. *Nat Immunol* **1**, 392-397.

Zheng, N. and Gierasch, L. M. (1996). Signal sequences: the same yet different. *Cell* **86**, 849-852.

Zohar, Y., Munoz-Cueto, J. A., Elizur, A. and Kah, O. (2010). Neuroendocrinology of reproduction in teleost fish. *Gen Comp Endocrinol* **165**, 438-455.

Zubrzycka, M. and Janecka, A. (2000). Substance P: transmitter of nociception (Minireview). *Endocr Regul* **34**, 195-201.

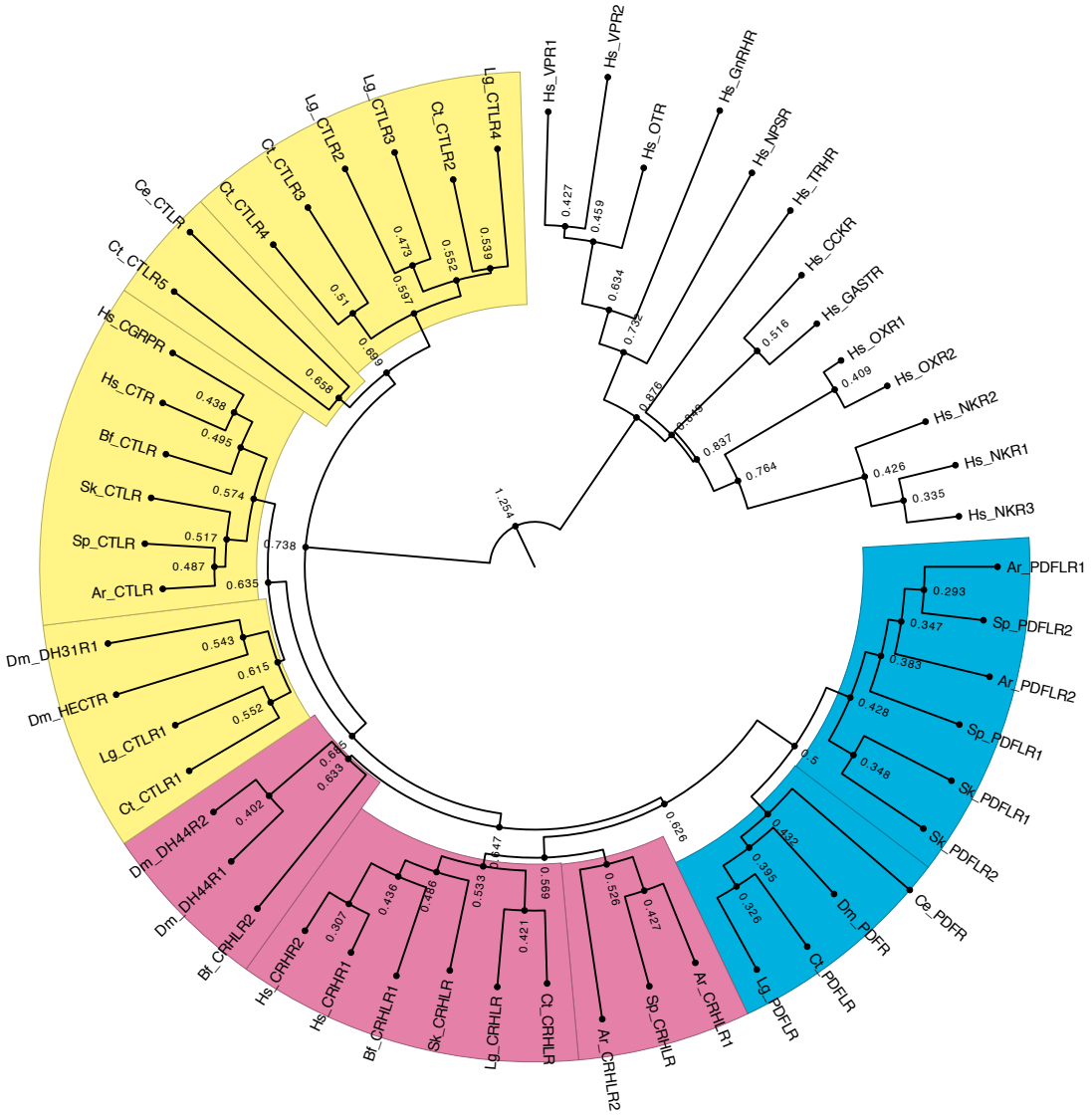


Fig. A1. Phylogenetic analysis of secretin-type receptors. Neighbour joining (NJ) tree (with bootstrap values out of 1000 represented in decimal format) of candidate secretin-type receptors for selected neuropeptides/peptide hormones identified in the starfish *A. rubens*. Calcitonin (CT)-like receptors are highlighted in yellow, corticotropin hormone (CRH)-like receptors are highlighted in pink and pigment dispersing hormone (PDF)-like receptors are highlighted in blue. Representative *H. sapiens* rhodopsin β-type neuropeptide receptors have been used as an outgroup. *Ar, Asterias rubens; Bf, Branchiostoma floridae; Ce, Caenorhabditis elegans; Ct, Capitella teleta; Dm, Drosophila melanogaster; Hs, Homo sapiens; Lg, Lottia gigantea; Sk, Saccoglossus kowalevskii; Sp, Strongylocentrotus purpuratus.* Note that sequences used for phylogenetic analysis can be found in Mirabeau and Joly, 2013.

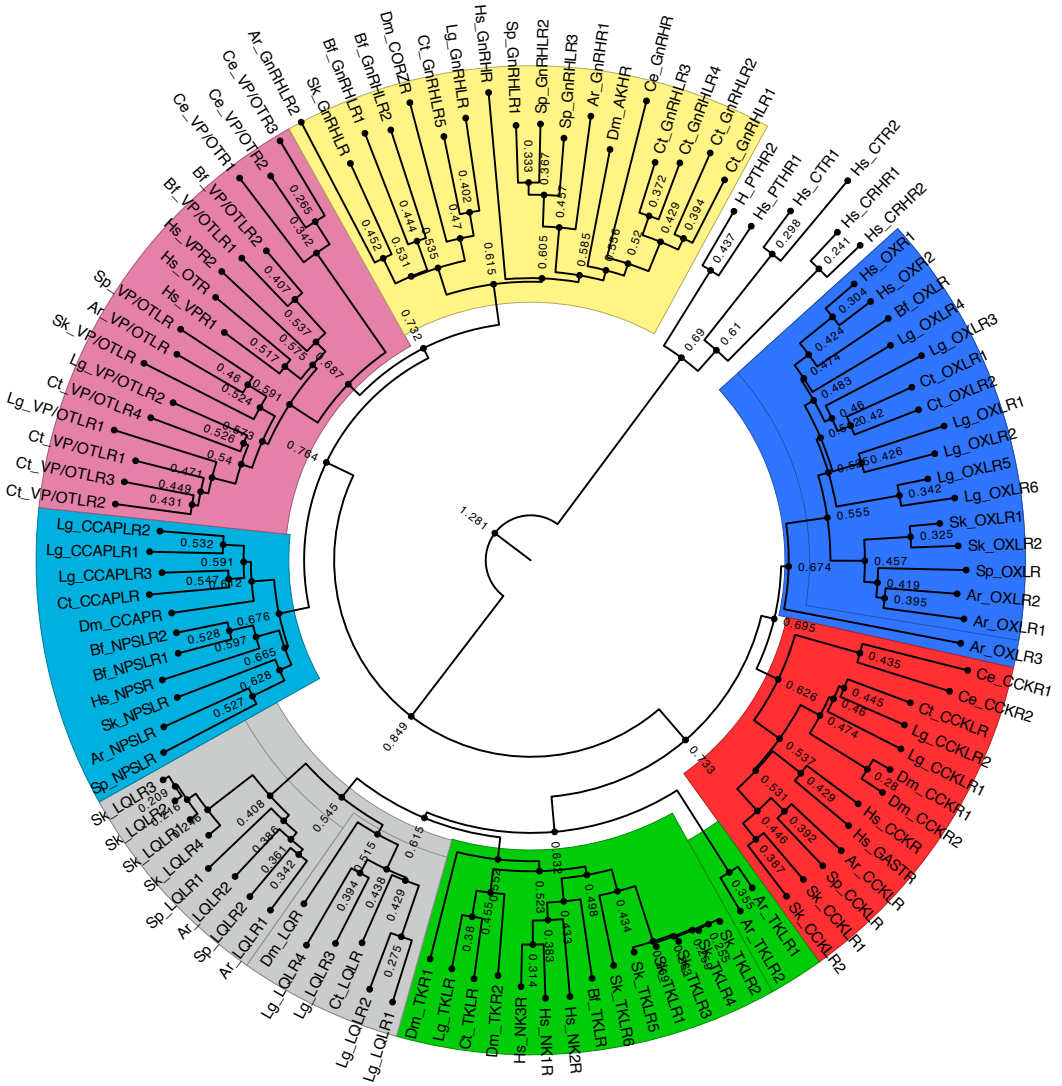


Fig. A2. Phylogenetic analysis of rhodopsin β-type receptors. Neighbour joining (NJ) tree (with bootstrap values out of 1000 represented in decimal format) of candidate rhodopsin β-type receptors for selected neuropeptides/peptide hormones identified in the starfish *A. rubens*. Gonadotropin-releasing hormone (GnRH)-type receptors are highlighted in yellow, vasopressin/oxytocin (VP/OT)-type receptors are highlighted in pink, neuropeptide-S (NPS)/crustacean cardioactive peptide (CCAP)-type receptors are highlighted in light blue, luqin (LQ)-like receptors are highlighted in grey, tachykinin (TK)-like receptors are highlighted in green, cholecystokinin (CCK)-like receptors are highlighted in red and orexin (OX)-like receptors are highlighted in dark blue. Representative *H. sapiens* secretin-type neuropeptide receptors have been used as an outgroup. *Ar, Asterias rubens; Bf, Branchiostoma floridae; Ce, Caenorhabditis elegans; Ct, Capitella teleta; Dm, Drosophila melanogaster; Hs, Homo sapiens; Lg, Lottia gigantea; Sk, Saccoglossus kowalevskii; Sp, Strongylocentrotus purpuratus.* Note that sequences used for phylogenetic analysis can be found in Mirabeau and Joly, 2013.

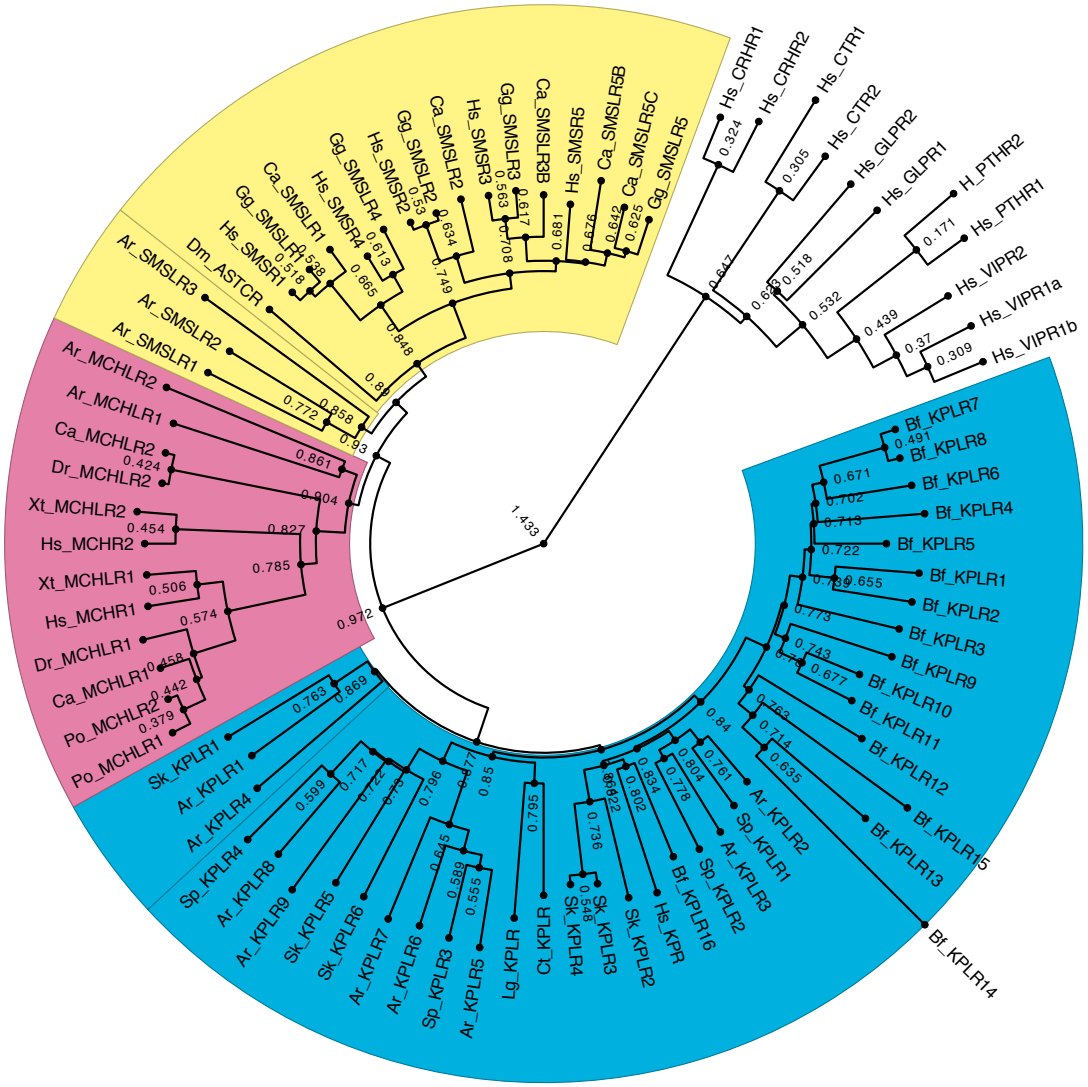


Fig. A3. Phylogenetic analysis of rhodopsin γ-type receptors. Neighbour joining (NJ) tree (with bootstrap values out of 1000 represented in decimal format) of candidate rhodopsin γ-type receptors for selected neuropeptides/peptide hormones identified in the starfish *A. rubens*. Somatostatin (SMS)-like receptors are highlighted in yellow, melanin concentrating hormone (MCH)-like receptors are highlighted in pink and kisspeptin (KP)-like receptors are highlighted in blue. Representative *H. sapiens* secretin-type neuropeptide receptors have been used as an outgroup. *Ar, Asterias rubens; Bf, Branchiostoma floridae; Ca, Carassius auratus; Ce, Caenorhabditis elegans; Ct, Capitella teleta; Dm, Drosophila melanogaster; Dr, Danio rerio; Gg, Gallus gallus; Hs, Homo sapiens; Lg, Lottia gigantea; Po, Panaeolus olivaceus; Sk, Saccoglossus kowalevskii; Sp, Strongylocentrotus purpuratus; Xt, Xenopus tropicalis.* Note that sequences used for phylogenetic analysis can be found in Jekely, 2013 and Mirabeau and Joly, 2013.

A. Partial (5') CT-like receptor (Ar_CTLR)

RFTPRGDHTSIYKKPVHTWLILVLVVLLNAGVISGDKMDTQTPQQRERHINESVLCE
YRMRTELPPDDEYVYCNMTFDTWDCWNYTRAGSTIVQPCPEWIPDSNPNKTASKICM
PNGQWFRHPETDDTWTNYTSCNYRVKEDLDSNVHIVYYFGYGLSVISLCIASFIFFY
FRSLGCPRVTIHKNLFISFILCGIVWLQWFALVTQNPALIADNPAFCKILLVLGNYF
NTCNYFWMLSEGLYLHTVIVVAVFSENHRLHLYYIIGWVLPMIPVTIYTSFMLVYHQ
GVCWLAESEYEWCVASFIILVLFVNLILLLNIVRVLVTKLRATPASGSNTYTRAVRA
TLILLPLLGLHYILFPVKPKGNKVAETFYNYFIAVLLSFQGFFVACIFCFFNGEV

B. CRH-like receptor 1 (Ar_CRHLR1)

MATHVLTVWITRTTYSLVLILLLASFSCTYGQANNTTTAIPTNTTDFLPPITSTEDT
NMSFTFIARTNASASPALPPEPQQYCNASIMDGLQQQFNREYMDYHKLTEISDEDTC
NLRKSVDEYLVDYCSPASAVYCDSYFDTIMCWPPSPVGLNSRPCPEELNGVSYDTSD
NVTRYCNLDGTWANGNKSNYSLCSPRNLPEMTGHDAVLKIMVYIGYSLSMGALLGAF
CIFLFFRSLRCVRNYIHWNLVSSFLLLYICFFTTTAATNTYSEKPGLHWLCRLSFTM
MMYIMMTNFFWMFVEGVYLYTLVVRALTVRKNRFWLYCIFGWGGPLPFVITFVILKL
FNSKDCWVDDAPEVYLIVAPIICVILVNLYFLVHIMCILFTKLRASHSLETQQYRKG
VKGTLFLLPLLGVTYLLFLLGPISVQVPQRPPSFYVYQYLNTILSSIQGFLVAIIYV
FLNQEVQNVIKRKIRRWREENTLPTRIVSRRGSNQRNSIGNIFVGRRSVNCNTMADD
DTNVSAFKLPACNRSPDVYSNGSKVPPFIRLDEIQPLTPASETHASPQTLYSSLGSN
GADSGDMQQVNLEEKEADASCNVTNGSTDTDCNELPLLAVNACGAQSPNSSPDQSTG
EELDRVIRKESQIATEKDQDLPYRTVDNNGLSNGNHNNNNSLSPRLARPKTSPFTGG
RTRVTFAEDPVQENLKRPACCDLGKNNEGLARSPTRKVEEPIARGHVEGDESLEGGY
SSKQDPHRSPLIPCIRKKPWATKPNGTIAFPKAPEGTPV

C. Partial (5') CRH-like receptor 2 (Ar_CRHLR2)

DPLLLVWLLAFIMNLGATGAGAQDEMNQEYYERVIHERAMDCVRRYQNVTVPGVDPN
FSCQFFFDGVLCWPPSAPDTKVEQICPLDRYEVALVQRTCFSNATWEQASPQSLYLS
DSCRCVHCLSLAVLNNLGYALSLAACAVAFLIFLCSRTLHCTRNRIHWHLISSFMLK
YLVYFAQIALFNTKELHLTRILDTLLFYFEQTNYFWMFVEGMYLHTLVVFALRIDAD
KIMFWVYCIIGWGLPMVFVGSWAIVSVALYPDSQFINMSQDAVIVHEYVCIIGPILL
VLVINCFFLLNIVRLLMSKLRMQGNRTSDIQHYRRAARATLVLVVLLGVGYILFICL
SFLPTDLPPLGNLILKYVNVLLASTQGLAVSVIYVFLNTE

D. PDF-like receptor 1 (Ar_PDFLR1)

MTTSEKECDFLKSMQTESEDACPIEYDKTMCWPETPLNTTAYQLCPDIFPQLDPTKF
AQAFCREDGTWGLFDYLPCFRKETEELDELRLNISDEKLEFTKDVIYGAKAMEIAGL
FLSFFCLVFALCVFSCFRSLQCHRTKIHKHLFAAYIARLFIEIILGVNQIYKLTQAD
VGFAPSVISSAPALCVILEVLREYTRLCTFAWMFIEGMYLNSLLSSAVFGKPKFLVF
YIVGWVVGIPLTITWGLVMHFINGSDRCWYDNITSGNCLLIIEIPRNAFLAMNLLFL
INILRVLVTKLRESNTSETKQVRKAIKATFVLLPLLGVANLIWLIPQAEASDSRAWL
IIYYYGILFLDAYQGFFVAVLYCFMNAEVGSVLRRKWSTWRTYHNPNRRGQNMVVTT
ATDVCTTEHAQFS

E. PDF-like receptor 2 (Ar_PDFLR2)

MVRSEDECDGIYQQFYAEERDALMCEPNWDTTLCWPATSANSTVTVRCPNFFWFNPD
NSLTRYCSPNATWETGGTYQDCFTEGAKSTINGNESFGNVTTEQIEFYIEVLEGANF
MRLIGVCLSWVALVVALLIFSSFRSLNCNRIKIHKHFFAAFLIRMTLEVILVADKRH
RDQSTMYNGVNDVKSIHLTPVLCELFETIREYGRLCAFFWMFIEGLYLTFLVSANVF
SKPRFRIYYCIGWLIPLPTVMAWAIAMYLTSKERCWKDHAVSNYYLIIEIPRNGALI
INVLFLVNIVRVLVTKLRESNSSETKQIRKAAKGAVLLIPLLGVANLVWFIPSPNVT
DSRASIVLYNYLFLFLDAYQGFFLCVMYCFLNHEVRMTIHRKWSTLRYYRDPYPFLR
RTSVLTSTSDMMSTTRAAYKPSQKMNGTAGGEYLKMAAIDEMSSTHA

Fig. A4. Secretin-type receptors identified in *A. rubens*. (A) Partial (5') calcitonin (CT)-like receptor (Ar_CTLR). **(B)** Corticotropin releasing hormone (CRH)-like receptor 1 (Ar_CRHLR1). **(C)** Partial (5') CRH-like receptor 2 (Ar_CRHLR2). **(D)** Pigment dispersing factor (PDF)-like receptor 1 (Ar_PDFLR1). **(E)** PDF-like receptor 2 (Ar_PDFLR2). Note that transmembrane domains were not predicted for partial G-protein coupled receptors (GPCRs); predictions on GPCRHMM (http://gpcrhmm.sbc.su.se) are restricted to a GPCR model and hence contain seven transmembrane domains (Wistrand et al., 2006).

A. CCK-like receptor (Ar_CCKLR)

MATATTAYPYSLIDSSLPPVNSTFLVTSIVDVNSTNSSLITEDFDDDRNRGVRIGFG
LNIYLTATLYGIVFVLAIVGNILVLVTLAQDKRMRTVTNMFLLSLAFSDLLFGIFCM
PFTVVGNMLGRFVFGAVICKIVPYIQGISVTVSVWTMVVISLERYHAICNPLSSRVW
QTKAHAYKAIVGVWMVALFLNLPAVIFSKLFSFNSGTVFRCDEIWPATLYRTIYRMC
LFVILMVAPLFTMLTAYGLIIRELRRGMKLEQCGADNEKRENGIAMKNMGDEASCSL
NEKKTKKSDKKPAQATMRSTSTSGAKKRVVKMLIVIVALFFVCWTPSWVGNIWIMIS
EKSASEHFGRAEVTIFKLMTYASACVNPIVYCFMNKRFRQGFLNAFSCGRRGRAGDR
ATASGDVSRFQSTRRTNVPRPSPTNYTNVSSDSSV

B. GnRH-like receptor 1 (Ar_GnRHLR1)

MATTSVNPTLITYIYETVEVNDTYSNFTKAGPTSDGNGTDYHSGISIYDDYLIRLIL
YVIIFVVSTIGNTAVLCSLIKGRRRKSRVNLLIMHLTVADLMITFFNIPTYFIWLIT
YQWYGGDIMCRSVMYIAMVGTYASPFILIVISLDRFASIVFPLSVRQADMRCKIMLR
VAWAACIIASIPQLLIHQVMSPKSDPDFTQCVDYGFRKNFPVFWNLYHWFVMAATYF
IPLTLIIGCYTSIVIKIFGNSTIRSYSGNNGNRMTLRRSGVDTLPKARVRALKMTGA
IVTAFIVCWTPTCIEGTITHANPALGEATPIWLNHIMLAFGFSNVCIDPIVYGMFTV
DFRRHFPGCFDCWGGRNILRGRRSGRSTTSSNNYPPVTYASSTRCGTTTGVSGNYHV
MELDSRHGTAEKCV

C. GnRH-like receptor 2 (Ar_GnRHLR2)

MSVQYTSDESEFALCDLLSSATECDENTTIALPEFTPFTAFKVGVLALMIVFSTVGN
CIAIAVTCKIRGRRKSTVTTLILNLAVSDLIVTYCHMLVHMIWYSTDAWLGGEALCK
IAKFFSNFGLFASSFITVDVGLDRCLAVLRPLGHRQRPFHIKMMIITSYMVAFLFSI
PQLIVFRLEHWPFDPEIDFWQCVTNVGIPNVYIAVYTTLVVLAQFVAPLGIMMVAYG
LIFMKVRQKIVMKETSENLSEMRQARSKLFLRAQRRTVRMAVAIFIVFAINWLPYAI
FGLWYVWFPGQTYNMYVFEVAFMFGLSNSCFNPLIYGACNVRYCHKICACFGIRSKP
GMDTSHSDRSNKTTMYSVRWQSSKGGNNTHKNNNKDNSWTKGTAAVPPERQQMITTT
T

D. LQ-like receptor 1 (Ar_LQLR1)

MNISYSNDSSGGYSDVCVEYYREPVWLQVLFGIIFGLITVFGIGGNAIVCYIVLGHR
RMRTVTNYFVVNLAVSDQLMAVMCVNFTFYSTLYMTWPFGPVMCKAVSFFQSVSVSV
SIFTLVAISMERYMAIIHPLRPRLGSTGTLLVIGFIWISSGALGLPTAIYTSVYVEN
GITYCSEEWDQRGNYSFATMVLQYFLPLAVLAMAYSRIGIRIWARKTPGEMEANRDR
RMTESKIRLVKMFAVVVFLFAICYLPIHTFNIVQDVYQYVLCYQYIRIIYVTVVAVA
MSNCMYNPFIYCWMNSKFRDGFKNVVRCLPGVRRYRDRQQANQSGSFRGLKRTSTLN
THMQS

E. LQ-like receptor 2 (Ar_LQLR2)

MSGLTISNFSETINSSDSSDYDYPTYEYYEEPVWIQAIFIFLYGTVSVIGIFGNGIV
CYIVLGHPRMRTVTNYFIVNLAIGDLLMAAMCVNFTVYATLYNQWPFGELMCKLVSF
SQTISVSVSIYTLVAIGVDRYFAIIHPLRPRMGSKETLFVIGVIWVISIALALPAAL
FTTLESQGGSTFCTDGAWEDSLVYSLICMVLQYFLPLAVLMGAYIRIGKRIWGRRTP
GEVEAERDKKMSESKARLVKMFATIVLLFALCYMPIQIYTIMQDSNRSILSFYYIKI
VYLCCLLSAMSNCVYNPFIYCWMNIKFRNGFRSVFRFLPCVNYQGDWNGFTGLRRAN
TAQTQTETMSMGSRNDRGRWWTQDQAKIAKNGRASNSDGETRTSFM

F. NPS/CCAP-type receptor ("NGFFYamide" receptor) (Ar_NPSLR)

MATIPAYDHLVTDSVMAGYSLNDTASTVMVPTGLPSTLEGAPNATTSVTYFSDGENR
LSFYGGFQLIVLWVLFGLTVIGNSTVLLAVYVIRHKKSRLNFFVAHLAASDLLVGIV
NNGYEVLYRYLGEFYGGMVFCKIIRFSQAYVINASSFQLVALSLDRFFAIVFPMDFS
GSGKRANLMAVTAWIAPLFASIPSAIVFEAAVDSWGKTHCIPPPLVPGSWQYKVYTL
YVVSGFFYIPLIIISTCYIFMVISIWRRSKYMMGQKPEKVKGKAAKNASKEKEPMKH
RASSRGLIPKAKIKTLKMTVSIIVAFIVCWCPFSVFYTLDAFGVIFIDEANLNTAFR
ASAFIQNLPFLNSAINPFIYGMFSTNICQELRRFSVINWMATKLRCCKSWRPSRYGR
STTLRTDTNLTDMSEGASGTHRGHTRPLYVYTPGRTSNSDHSRDLSNSAM

G. OX-like receptor 1 (Ar_OXLR1)

MEPEMDMENPAFNSTLYSSMQLDSEYNQYILDFLKTKVFPAPYEYLIIAIYGLVFLL
AIIGNTLVCIAVLRNEHMRTVTNYYIVNLAVADILVSLVCLPVTVVVDVSETWFFGG
VLCLIIPYIQNTSVCVSVFTLTMIAVDRFLAICYPLKFQSSARRTIISVIVIWCVSL
AITLANPLTMEVTSQQWMDGANPVWLSKCVEKKWDKTEGRKAYYVVMLVFTYIIPLY
IIGLAYMMVCVRLWSGIPSDETSCQGKKAGNNNSGHAKGKRVMNKSAEAQLQSRRKV
ARMLIVVVVIFAVCYLPIRVLNVLSAFGAFSQANVDMMAHSANLGVAYLIAHMMAFI
NSAINPLIYNFMSAKFRQAFKMTFNCCGSDSKRHSRYTQRYTKCQSFNQSSAISDSR
ANTNTECFSMANRPNSIVHTTEA

H. Partial (5') OX-like receptor 2 (Ar_OXLR2)

PPRWDEFAPSALNPTGPPETDDYVDYIYDFIKGMVYPDAHEWFLIAAYVVIFVTGVI
GNFLVCFAVLRNEHMRTVTNYYIVNLALSDILVVLVCLPPNVMVDATETWFFGDLAC
HIVPYIQEVCLSVSVYTLAAIAVDRYLAICRPLKFHIRASRTLFIVAAIWVTSFTIM
APQPVFTRSEPSSHAGYIGKPIWLMKCYEKAWKGTIWQKVYHISKVVVSYVIPLLIM
AVAYARVCLRLWSGIPTDESHRPALGGRSAGQTKPTMNKSTETQLESRRKVARMLIV
VVVMFAICKLPTHVLNTIRYMDGFNNMKASENTESSGRVALHVSFLIAHFMAYLNNV
INPVIYNFLSARFRKEFRHVFSCCPCVTAPDREGKRPDSTSYHRCRDFAARSSNNGV
TSQSDFRSHDHNTEYLQLSKLGKSSGTNNNSNIVINTKPLAPNPES

I. OX-like receptor 3 (Ar_OXLR3)

METTTLTSVLYPNYTASVFPNSDYIMDGSDADVNGTSIQPVLTPGVLAVTTIAYGII
FFFGVTGNALVAIVIWQNADMRSSTNYFLVNLSIADLMVILVCLPPSMLELYFPDWM
LGKFMCYLVTILDSATAHASILTLLAIAVERYYVICMPFKANYTCTSKRTLIICCLV
WLIAFISSVPLALTVVYSNTEAGCVCGSFANTTLKEAYVGGIAVGFFLLPCAFMAII
YCVIANTLRKHDNYMASIRNKESLGPASGTPVYREGRLVELATMPRKNSGDSMVENN
YVPLKNHVTSNSTNEEHEPTVAFIENEANHEAAMRPSNQSSSNSTCQNLANAATRQA
HRRVVLMMASVVAVFFLCWFPMRTVILWQIFAPKESIRAWLGANMHTFQVMIAIFRI
LVYINSAINPILYNLISSKFRVAFRSVIRCENARRARSKRSISTRSFTSTSSLRIHS
RTTDYHG

J. TK-like receptor 1 (Ar_TKLR1)

MVDYKDINSLNFTFVHNLDQESRILWSVFCSIIMAVAAGGNLIVIWIVATNPRMRTV
TNYFLLNLAVADALIATLSMPFMFSYIVTQNWALGIGMCKVVRFFGLVSTSASVLSL
VAISIDRYRAIVHPLLPRLSKSYIVCMIIFIWGGSSIFASPLLVYSTVLTFMYNDGV
IKNQCFIKWPDGIYRRIDFTYNVASFAVLYCLPLTVLAGCYTVIGVKLWRGDVLGEY
IPNRARQLKAKRKVVKMIIVVIAVFAICWFPLHVYQFLGYLHEEVYTQSYAVHIYST
IWTLAMSSSMYNPFIYCWLNDRFRAGFKRVFHCLIPGSKKQRDNKGIDRRTPLTSTT
SNANSLMTASTTGMHRQVPNGKPAILCRTSLIGISVVDSSEDIF

K. Partial (5') TK-like receptor 2 (Ar_TKLR2)

TVSTVASVLSLVAISIDRYRAIVHPLLPRLSKTCIVCMIIFIWIGAISFASPQFFYS
QLITFGYSDGDITQCYIIWPDGIQREYEFWYNVTSFIVVYCVPLSVLAMCYTVIGVK
LWRSGVLGEYIPNRAKHIKAKRKVVKMIIVVIAVFATCWLPLHIYQFLSFMSYSVYQ
KSYSLHIYLSVWTLAMSSSMYNPFIYCWLNDRFRAGFKRVFRCSQGTSDTRHSQRQS
GTPMSPTSRMTSSTGYQRTTNGKQELGCKNSLLSASVVDSMEDTC

L. VP/OT-type receptor ("asterotocin" receptor) (Ar_VP/OTLR)

MTPSQVLTTRVPVVATDTKESNISQEGAATYGLMTTSTPYHPTRDNQLAIVEILITA
SIFVLAIVGNTIVIAVLWRRRKSLSRMHYFIIHLCVADLTVAFLYTLPQMLWDITYK
FYAPDVVCRLVKYFQLFPVYLSTYILVMTAIDRYLAICHPLMGLRRNQTFRMRVMVL
IAWGIAVVCSLPQLAVFKLKNRTPNSHQEWMDCRANFQNSAGVKAYITFFTMAIYVI
PSLILAAMYGMICLTVWKNMGKNYKSSDTKTPQQPAPQKENNNWEDDETDKLNPDHV
EMQQPTVTKKESSQKTQVQYRRHGAAGKVSRAKIKTVKMTLTIVTVYVVTWAPFFVA
QMLSVWTTVQFDSLFFAIALLLTSLTSCINPWIYFAFSSNAGRDIKQTFGCLKVSRQ
DNGLASDKTHHDETSSRFVSTTGTASAHM

Fig. A5. Rhodopsin β-type receptors identified in *A. rubens*. (A) Cholecystokinin (CCK)-like receptor (Ar_CCKLR). **(B)** Gonadotropin-releasing hormone (GnRH)-like receptor 1 (Ar_GnRHLR1). **(C)** GnRH-like receptor 2 (Ar_GnRHLR2). **(D)** Luqin (LQ)-like receptor 1 (Ar_LQLR1). **(E)** LQ-like receptor 2 (Ar_LQLR2). **(F)** Neuropeptide-S/crustacean cardioactive peptide (NPS/CCAP)-type receptor ("NGFFYamide" receptor) (Ar_NPSLR). **(G)** Orexin (OX)-like receptor 1 (Ar_OXLR1). **(H)** Partial (5') OX-like receptor 2 (Ar_OXLR2). **(I)** OX-like receptor 3 (Ar_OXLR3). **(J)** Tachykinin (TK)-like receptor 1 (Ar_TKLR1). **(K)** Partial (5') TK-like receptor 2 (Ar_TKLR2). **(L)** Vasopressin/oxytocin (VP/OT)-type receptor ("asterotocin" receptor) (Ar_VP/OTLR). Note that transmembrane domains were not predicted for partial G-protein coupled receptors (GPCRs); predictions on GPCRHMM (http://gpcrhmm.sbc.su.se) are restricted to a GPCR model and hence contain seven transmembrane domains (Wistrand et al., 2006).

A. Partial (3') KP-like receptor 1 (Ar_KPLR1)

MTSYEVDITGVLNDTSSKSWLDQAMEPGPETVFVPVVWSLIIAVGVTTNAVVVYVVM
LHMKMTTVTNYYIVNLALTDISYLLFCTPFTTLTFTLYGYLTEGMCRVISYVQQASV
HATCMTLTIMSADRYFAIVYPIKSMKYRTRRLSFLINLSVWILSYIMAIPAPIFVEA
ESHFYYTGEKYFCFEAFSTPTKQAVYYTYVLVFTYILPLIVICACYSLLLRTTWTRH
HPSAQNSHRTRQNLLQKRRMTRMVLVVVCL

B. KP-like receptor 2 (Ar_KPLR2)

MTTSITQTIGFLVNSSECCHESIAFTTKTLFGETTTTSSIPTITTDEAAVVQATAGQ
AAWLIPVIFGLITVCGVVGNFLVIYVIVRHGHMKTVTNYYIVNLAVTDISFLLCCAP
FTATLFVSPNWLFGRFMCKFVFYMMQVTGQCTCLTLTAMSVDRYQAIVHPIKSLKSR
TTRVACIVNMCIWIGSLLISVPVAIFFDLSLYINLLVCSEMWPLDIMFPGYAVFSFI
LLYVIPIFTISVCYSLMLRKLWSRVSPGEDNNSHLNNARQKRKVTRMVLVVVLVFAI
CWLPTYIINLWIRLDPYFPKTNATYIFKMAAHTLSYANSSVNPFVYAFMGENYRRYF
KKAFPVCFRQRVRRPTDFTTSRTDPTGNFNGGTAGIGGRTVETVALD

C. Partial (5') KP-like receptor 3 (Ar_KPLR3)

VPRRKYSKMSTPAGLLNVTTAIPNVTEASMDDGGASVYTKLVPALIGIITLVGLVGN
SIVVYVIVCQGHLKTVTNYYIVNLAITDIFFLVFCAPFTASIYATPSWLFGRFMCKF
VFYMMQATAQATCATLTAMSIDRYYAITDPLKALKTRTPRVAIVVSVGIWTFSAVLA
IPVAVFFDIDVVQFQNQTYDICDEMWPLKIVNQGYGVYCFVMLYLIPLTIIVVCYSI
VLNRLWKAVSPTEETHAPVHLRMLIQKRRITRMILAVIVAFAACWIGTHIMSLWRRL
DVNFPRTSTSALVFQTIAHLLMYFNSCVNPFVYAFMGGNFRKQMARAFPFLAGKKSA
MTEGPTGSLIKSKSTQV

D. KP-like receptor 4 (Ar_KPLR4)

MTTTVATIFPLSTLGTNNDSVGNETFYQQPADLKPPHHALLMPILFGILMVVGIVGN
TLVIFVILKMRQFKTVTNYYVVNLAVADILFLCICAPSTAAQYGSPSFLGGRFMCKM
VYYMQSVSAQVTCLFLVAMSIDRFQAIVRPLKSLKTRTLHNAATISIVIWLFAVTVY
IPLLVFFDTVDIPYNGGIILLCKEFWSKTWSELFSIWIFLFTYTLPLVVISICYSMM
ILNLWQRVVPTDALSGPANDRNLRQKRKITWMVLTVVIVFAVCWLPVHVIQIWMEFD
VNFPYTMATLDFKAAGHALIYINSCANPFIYTFLGENFQKNFKKTFHCCFKKVRPTR
EGAGHVANTGTQQGSSSEQAGTRRKTGGGRTQMSTLTSSNVEAKGIEQIDNNHC

E. KP-like receptor 5 (Ar_KPLR5)

MMSTADHQLGLLHQIQISNLSTSIMEDFEVENYYDSNNTSDYVFDISVSIFNATSRI
LVPTTFFLLALMGLVGNCSVMFIICRHTDMQTVTNYFIANLAITDVATVLFCILPTA
LQNSGIIPMSTGVCKGVNYIQFVTVQATCCTLTAMSIDRYFLIVHAVRSRRSRTTNK
VLIINVTIWAVSFMMHSPVAVVSKVTSYNYCETVFGTVRGERVFQTFATLSMYVVPL
IINLVCYISILLQVWTRTARGTESAQAQERAVRRKRKITRMVFVVVLLFAVCWAPKH
FFRMWIAFDYVEFHMTSNRHYVLMGSLQFVALCLAYGNSCVNPFVYAFTTTSFKKYF
KKVFKPCCRFEDRQARTSVNQSRVSKVITGEESLV

F. KP-like receptor 6 (Ar_KPLR6)

MTDSQAFVTDIATTMVEVYGPTNLTSLFDSTSSYISGPCTEGNSSLGDGVTPSGMTP
GLQFPDSFLNPIRIVMPIIMAIISIIGLVGNGTVLCIIFKYREMQNITNYFIANLAL
TDVAMLGICAIPTAAGLAGMKMEEGLCKGVNYMQFVAVQATCCTLTAMSIDRYLLIV
HAVKSRKTRTTTRAIIVNVAVWIASFIVHSPVAIFYQLTDQGCELNIGKDGNGAKYY
YLGAFLSMYVIPLMIILMCYAKILIIVWRKTSAGTESAQAHERSIRQKRKITRMVFI
VVLLFAVCWAPIHCILLWEQFKTEDITNLSTAMAFACLRLFALCLAYANSLTNPFIY
AFTTASFKKHFKKVLSCSTPTESKEQKTMGPTSTYKKIGKYVNEYSSVNTCDTKV

G. KP-like receptor 7 (Ar_KPLR7)

MEGTCGNCSSDVGGDVDPITGIIFPVLMVLIAAVGITGNSLVLHIIFRHRDMRTVTN
SFVASLALSDIAMLTMCVIPTATVAITETWKLGDFLCKAASYTSFVTVQATCLTLTA
MTVDRYYLIVHAVKSRNTRTVMKAVVINVAIWIFSSLIHLATPVFTAVDSDSNCANT
FPNPDVDAKIYGVYAFLGMYIIPLMLIMFCYAKILIQIWQKTSGGTESAQAHTRALK
RKRKITRMVLIVVVLFAFCWAPLQIFIIWTNFNYAQVSEMRKDVFLFLRGSFQCMAY
ANSCVNPFVYAFTTTSFRKYFRKMFATCRGNVYKDRTSISISMTTKTERLAAEEDSS
I

H. KP-like receptor 8 (Ar_KPLR8)

MSSGDFLSDSYSMYYNDNATVEAASLGPSAVLVPLTLGIIAVLAVVGNALVIYIVLR
YRNMRSSVTNFYIMNVAISDIVFVVICVPLTSVSYGMTYWPFGQFFCKLNAYMQCVS
VQATCTTLTAMTVDRYYVIMSPLASRRTRTICRAGMVCASIWIFSAVVHIPVAVFFK
IEVINWFGEINEYCKFTLHKPAALSGYFIYLSLSTFLLPLTIIAVCYSLILSHLWSL
SRVGRCREPTSTAGYDPLSQQSRSGALPPAQSSARTASKTWKTTRIVLCVVMLFAVC
WAPLQAFNVWNAIDPEHHSSSVHVNNLRVFCLCLAYANSCINPIVYALAGTSYRHHL
HQMVSGSNKSGRINSRFSPTGGRNSSYRSVRRMSRTEATSVSTCV

I. KP-like receptor 9 (Ar_KPLR9)

MEYPDVSYSYSYSSFYYSSNSNFTNGFEGETGVHAILVPIIFGIITVVGLIGNGCVV
IVIARNRCMRTVTNFFIMNNAITDMVFVVICAPVTASQFILTDWIFGDFICKLVVFM
QYVSVQASCSTIMAMTIDRYMVILHPMRSLHARTIRRTTCINIVIWLTSFLLHVPVA
IYYEQVSEPNNGYFCGTNFVNITASKIYHFYAVIVLYVFPFLVMTFCYLRILRKVWS
KYLIVTSSTQTQRKRRWKITRMTLLVVILFGICWGPIHAVHLVALFKTATSATEIDW
SYYNFSIFCLCISYSNSAINPFVYAFSGRSYRSLLLSCLKRKPGKNPVSRMCSSRTK
AGSAQYSLEMNNLIRTNGQNGRSPKTRYTNQYIGGGRHIVRAVPE

J. MCH-like receptor 1 (Ar_MCHLR1)

MEGNLHSDLSTQVSSAKMYSLTTQHPVYQGENNTDYWDNSSSDAISNATENIQSVLF
QRFLAPLIFGCISIVGIVANGLVIIVLLKYANMRTIPNVYILSLALMDFMFVLSLPM
VGYQFCTNNWPFGYFMCKLVPGIDGFNQFGSVNIMTMMSADRYVAVVYPLSSMRYRT
KKTARIVCGVVFLISFILCTPSWYFMELKDGGYGLTYCIAKGPSWDKEGTLYTLYYA
CFGFLIPLAIIIVCYSTLLFKILGSKLRIRSDGTGTAQRASKRVFILTVSVIVVFVV
CWLPYNVVQLFHQFTTAEVSHAFSIAYAVSSLWCYTNSCFNPIVYTFIGENFKKNVM
HLCPFCTSPDDTKLDRYRSNRESSTYIRGGTLRLHSVQHEAPASPPAMTTNATAITD
ATSFTEAPANTQTYTFSNNNTHNSCGKE

K. MCH-like receptor 2 (Ar_MCHLR2)

MDKSTSIMLTTAGYSDILTPMFCANCTTPAGHPQDNSFLMSLQQISFIVTCILGTVG
LLCNGFIILILMRFPNMKTLANYFILNLALADFLFMISFLFLGHQLRVNHWVFGKFM
CRLIVPYDAMTQFLVIYFVMIMSIDRYFAICLPIKSMNFRTLRNGKIVCAMIWGVAI
LTTLPLWIYTEHTCVNGWCTCLARVSSNDEDDPRWWIIYTIIIGFCVPLTIVCICYL
LILQNLLTSNMQDSKNTLRRAARRVAILVICIIIVFILCFLPFYVFQLVYSTIDAEE
DASNALILIYIITWTLMYSHSIINPIVYALVGENFRKNITFMFCKRSSRHVMYTRQS
SMRTSSFRRTRNSLKSCNSNDPNGMRPHNGYYVHLDTGQHRGAREDPV

L. Partial (3') SMS-like receptor 1 (Ar_SMSLR1)

MDCMDTASDSFSTNCSLENCTEEETTKSVTQIRDWVAGIFLLIVCITGLIGNSMVIF
VVLRYAKMKTVTNLYILNLAIADDLFMIALVFLSIATLAGNNWIFGPVLCHVVFAID
GLNQFTSVFCLTAMSMDRYVAVCHARRVRGYRTLRLAICVNFGVWFLALAAASPMNI
VTRYTVYNGTPLCYFNFESVFGKDAVVITQKLFMIYTTLMGLVIPLVIISVCYSSIV
VRLKKLGTRTGKMKKSRKVNRLVFFVVLAFFMCWWPFYVWRTIVVFVPCLGQ

M. Partial (5') SMS-like receptor 2 (Ar_SMSLR2)

EEEEEEGNFSDIQLAGTIIGSLYMIITIVGILGNGTVIYVVLRFAKMKTVTNCYILN
LAVADAVFVTFLTLMAVSNFMETYWIFGAFLCKVYFGIDMFNMVISVWCLTAMSVDR
YVAVCHAMKSRSFRNLPIATAINASTWLLSILAAIPFVYFAKLEPGSGNYDVCYLYF
EPDRVKTSSQIMAMCVFIFNFVIPLTVIIVCYASITMRLRDMNKKTGKPEKSRKVNR
LVLIVVITFVICWAPFYIVKILFVFVDTFRRWNKTWILSDITMSLTYINSCANPFLY
AFFSDNFRKSFRKAWLCHSRNQAEVSQTSYASRWKFGKSGGKGKFNKKGKRFRLHDD
EGDDTNVSGYNNQYPLTATAVTSVVSESSYPRNGEMSQVNLDKQAVAKA

N. SMS-like receptor 3 (Ar_SMSLR3)

MASVGNSTEEPFADLYGDSDWIYSVTGSFSYIVCVIGLLANTWVICVLLCKLGLKTA
GNMYILNLAIADDLFLAGLALQAAYQITEVWHFGEFLCRAALAFDGLNMYASAFFVT
ALSIERYMAVSRSSRARLHRKRRQAAAVSVIIWVISILAAIPTLLASYYITNTREDY
ICITTFSNFGEPGVEFWNQAFITYNFVLGLCVPLLVSCLCYGMLIVQMRQVSMRNDA
GADVKRVTRIVTGVVVVFFLCWAPFYLVRMILVYNPSLMYWSGKPFVFELSLCFTYI
NSCVNPFIYALISEKFRENLPCLALRGRGYISGKRTAVCDRNSTALTDIHRNSVITN
NDSS

Fig. A6. Rhodopsin γ-type receptors identified in *A. rubens*. **(A)** Partial (3') kisspeptin (KP)-like receptor 1 (Ar_KPLR1). **(B)** KP-like receptor 2 (Ar_KPLR2). **(C)** Partial (5') KP-like receptor 3 (Ar_KPLR3). **(D)** KP-like receptor 4 (Ar_KPLR4). **(E)** KP-like receptor 5 (Ar_KPLR5). **(F)** KP-like receptor 6 (Ar_KPLR6). **(G)** KP-like receptor 7 (Ar_KPLR7). **(H)** KP-like receptor 8 (Ar_KPLR8). **(I)** KP-like receptor 9 (Ar_KPLR9). **(J)** Melanin concentrating hormone (MCH)-like receptor 1 (Ar_MCHLR1). **(K)** MCH-like receptor 2 (Ar_MCHLR2). **(L)** Partial (3') somatostatin (SMS)-like receptor 1 (Ar_SMSLR1). **(M)** Partial (5') SMS-like receptor 2 (Ar_SMSLR2). **(N)** SMS-like receptor 3 (Ar_SMSLR3). Note that transmembrane domains were not predicted for partial G-protein coupled receptors (GPCRs); predictions on GPCRHMM (http://gpcrhmm.sbc.su.se) are restricted to a GPCR model and hence contain seven transmembrane domains (Wistrand et al., 2006).

A. Partial (5') GPA2/5-type receptor

PPLFRFQSISTAKLHYFSHCCAFYYNTTENTHMAKMHNITEAVEDVDCPTSTSQPVR
SASPFEFLDNREASIPRGSREILDEYIQSAVVPPNKESSMGHSFFIPVNNESLMAVT
GEPFIGPFGADDTSDMGFREEVVDPHVTHVIYGSTCHDIVPKNYSKVNCSPLPDPFN
PCSDVMGYPFLRVIVWFIASTAMIGNLVVMVVLIAYRHKMTVPKFLMCNLAFADFCM
GIYLLMVAAVDVHTLGAYFNYAIDWQFGAGCKIAGFISVFSSELSVFTLTVLTIERW
YTIIYAIDLNKRIRLRQAGRILLGGWFFAIVIASLPFLGVAGYGVTSMCLPFYVRKT
DSAAAVAYVIFILLFNALAFIVICACYIKMYVTVRNPHSAMQRKDSKVAKRMAVLIF
TDLACWAPIAINGLSSIFNGKNALLTTDQTKVFIVLFYPINSCANPFLYAIFTRAFR
RDLFLLLAKHGMCEKRAMKYRPTYTSSPRSMSQVNSTTMRDVCPKGSSHRGSSASVM
TQITSDARPSMVSFDKESPDNSPKMNIELQFVVTPGSSLQSQIPEDTSISNQLEVSH
SPNTGSKAGGTKKLNEYSYILDGLPPVPEEHNSDRPEPSSKPFQTELDDNLNPCRPG
NPLIIEHCVSDNFVDMEPVNRESSSLLSYVAGNGDVDAALTNGNRTDIGIQSVTPND
HEVNSKLIQMNLTKADDVETTL

B. Partial (3') bursicon-type receptor

MELILMLFWAAVYGVMLPVFPATPTKFNCHRNSSLCSSVCCCYRPKNPNGGSVDIIA
VDCTRRNLTEVPLDLPANTVEINLSFNNLTALPADAFSHVTRLRHLKLVGNRIRTLN
EDVFRGLKELHSLDLQLNIFQKLPSKAFHPDLANLRTLNLNSNYIEDFPGHALLNLH
SLRHLYVEGNRLNRIPTEAISLVPELTVLLLSSNAITMVQDHAFQNNTKLIELSLNN
NSIQVISTHAFSGLYNLKGLDFRYNYHRIKNIKQGAFRNLTNLNALAISDVDNLSEF
PDLSGTTSLERLTIDRCSLTAIPTTFCDGMHHLKELDLHNNKIESIPSFSNCSQLQY
LNLGLNKLSSLESTPFQGLNQLSHLYLGENNIERIRADSFEGLSAAYELDLGNNDII
EIDEKAFLPLKQLETLRLNGNRFPDFPTAGLERLLNIYTFDNIHLESFPKPSELPKI
IIIISSYAYHCCSYISLNEVNSVKNTPEIKEIYTWTLSSSTYDYNGGNGSSSDWAEE
WYDSTWEPQYNLSGVPIQIHSNISCSPVPGPFLPCVDLFGSWFLRIGVYFVFLLAVI
GNAIVIFVILVSHTKMDVPRFLICNLAFADFFLGVYLGFLAGVDTSTLGVFHKYGAQ
WQLSPGCSLAGFLAVFSSEFSIYTLSVITLERFYAIKHALHLEKRMKLPHAVAVMAF
GWIFAVVVAIMPLVHISSYYKFPVCLPLDITTTIGKVYGSSILLLNVTAFVIIMLCY
ANIYMAIQGSHAWNCNDSRVARRMSLLVFTDFACWAPIAFFSLTSAFGLQLISMEGA
KVLTIFVLPLNSCANPFLYTILTKQFKKDCQMILRSLRNQVLRPRKSTNPSLGRHPS
MRSTQAHQISQLPWRTNGSISQDITGDGCMVSSFKTENGEPSSNVYTMSTMAETHTP
QQNGSCSNDHSHSDSSAMLDPMLLTNSDLDDLRRGSTSPLLDLMTQVTSKDPEAKPK
SKVRRKFSIPFPFFSTRSPTKEPTHSSSANTSIRRSPTPQRQKRSHSITEGLLHKMN
RKRPLTSMKSAPGISSPISRSSEKLDQGINGISIGICITVPDKQNQNKNVKQTEADE
KVYSIDVPKEPPKPPLSNFPIDQAGMETQTLLEIGASNNYFSENRSPHDTDSPGMLA
AHENPVRDQLSSATSSPQLHRPEKLVLVDSVVPADKDVPDNDSGLGTASSPLPDSTE
SMRFLPVSLETPEVVCNSELDDHSSRVTSQPSVLGGLFGH

Fig. A7. Glycoprotein hormone-type receptors. (**A)** Partial (5') glycoprotein α-2/β-5 (GPA2/5)-type receptor. **(B)** Partial (3') bursicon-type receptor. Note that candidate GPA2/5 and bursicon receptors belong to the rhodopsin-δ family of receptors (Fredriksson et al., 2003) so were not included in neighbour joining (NJ) trees. Note that transmembrane domains were not predicted for partial G-protein coupled receptors (GPCRs); predictions on GPCRHMM (http://gpcrhmm.sbc.su.se) are restricted to a GPCR model and hence contain seven transmembrane domains (Wistrand et al., 2006).